Journal of Neuroscience Methods 213 (2013) 43–62

Contents lists available at SciVerse ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

A procedure for testing across-condition rhythmic spike-field association change

Kyle Q. Lepage[a,*], Georgia G. Gregoriou[b,c], Mark A. Kramer[a], Mikio Aoi[a], Stephen J. Gotts[d], Uri T. Eden[a], Robert Desimone[e]

[a] Boston University, Department of Mathematics & Statistics, Boston, MA, USA
[b] University of Crete, Faculty of Medicine, Department of Basic Sciences, FORTH, Crete, Greece
[c] Institute of Applied and Computational Mathematics, FORTH, Crete, Greece
[d] Laboratory of Brain & Cognition, National Institute of Mental Health, Bethesda, MD, USA
[e] McGovern Institute, Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology, Boston, MA, USA

HIGHLIGHTS

- A statistical methodology is introduced capable of studying changes in the coupling between rhythmic local field potential (LFP) and neural spiking times.
- The methodology successfully deals with a problematic confounding factor present in more standard analyses based upon spike-field coherence.
- The method is capable of studying both per-frequency modulatory effects as well as the tendency of spiking to occur at a specific phase of a sinusoidal (LFP) rhythm.
- The method is effective both in simulation and when analyzing real data.

ARTICLE INFO

Article history:
Received 11 August 2012
Received in revised form 19 October 2012
Accepted 19 October 2012

Keywords:
Oscillations
Rhythms
Local field potential
Spikes
Statistical association
Hypothesis testing
coherence
Spike-field coherence
Generalized linear model
Point process

ABSTRACT

Many experiments in neuroscience have compared the strength of association between neural spike trains and rhythms present in local field potential (LFP) recordings. The measure employed in these comparisons, "spike-field coherence", is a frequency dependent measure of linear association, and is shown to depend on overall neural activity (Lepage et al., 2011). Dependence upon overall neural activity, that is, dependence upon the total number of spikes, renders comparison of spike-field coherence across experimental context difficult. In this paper, an inferential procedure based upon a generalized linear model is shown to be capable of separating the effects of overall neural activity from spike train-LFP oscillatory coupling. This separation provides a means to compare the strength of oscillatory association between spike train-LFP pairs independent of differences in spike counts.

Following a review of the generalized linear modelling framework of point process neural activity a specific class of generalized linear models are introduced. This model class, using either a piece-wise constant link function, or an exponential function to relate an LFP rhythm to neural response, is used to develop hypothesis tests capable of detecting changes in spike train-LFP oscillatory coupling. The performance of these tests is validated, both in simulation and on real data. The proposed method of inference provides a principled statistical procedure by which across-context change in spike train-LFP rhythmic association can be directly inferred that explicitly handles between-condition differences in total spike count.

1. Introduction

Many experiments in neuroscience (Fries et al., 2001, 2008; Womelsdorf et al., 2006; Witham et al., 2007; Pesaran et al., 2008; Gregoriou et al., 2009; Jutras et al., 2009; Chalk et al., 2010) have compared the strength of association between the times at which neurons fire and rhythms present in local field potential (LFP) recordings. A measure of association employed in these studies is the "spike-field coherence", a frequency dependent measure of linear association between a point process and a continuous valued LFP signal. Spike-field coherence is shown to respond to overall neural spiking activity (Lepage et al., 2011), making comparison between two pairs of spike-field time series difficult when the average spike-rate differs in the two spike-field pairs.

Existing approaches to dealing with this confound include the employment of neural rate-free measures of association and

* Corresponding author. Tel.: +1 6173531841.
E-mail addresses: lepage@math.bu.edu (K.Q. Lepage), gregoriou@med.uoc.gr (G.G. Gregoriou), mak@bu.edu (M.A. Kramer), mcaoi@bu.edu (M. Aoi), gottss@mail.nih.gov (S.J. Gotts), tzvi@bu.edu (U.T. Eden), desimone@mit.edu (R. Desimone).

0165-0270
http://dx.doi.org/10.1016/j.jneumeth.2012.10.010

transformation based techniques, such as neural spike thinning (Mitchell et al., 2009; Gregoriou et al., 2009), a procedure where the overall neural rates are made equal by randomly removing spikes. In this paper, an across-condition test is introduced based upon parametric modelling of the effect of the association of rhythms in field-type time series upon the intensity of the spiking process. By explicitly modelling the dependence of neural spiking activity upon both random, "background", influence and upon field-type rhythmic influence, the relative separation of these effects is made possible.

After a discussion of relevant background in Section 2, an hypothesis test directly comparing the nature of the association between pairs of neural spike times and LFP rhythms is introduced in Section 3. This test, based upon the generalized linear statistical neural modeling framework presented in Truccolo et al. (2005), uses all available data and accounts for differences in firing rate within the maximum likelihood statistical framework. The test provides a statistically principled approach to inferring differences in association across spike-LFP time-series pairs with different spiking rates. In Section 4, the test is demonstrated in simulation and in Section 5, the test is performed on real data. The paper ends with a discussion in Section 6.

2. Background

Coherence, analogous to cross-correlation in the time domain, is a theoretical quantity linking two time series in the frequency domain. Non-parametric coherence estimators are common and have been successfully employed in diverse sciences. In neuroscience, background material on field-field coherence (coherence between two field-type time-series) and spike-field coherence includes: Brillinger (1975), Brillinger (2001), Rosenberg et al. (1998), Halliday et al. (1995), Amjad et al. (1997), Jarvis and Mitra (2001), Mitra and Bokil (2008), and Lepage et al. (2011). Coherence in the neuroscience setting has been used to characterize neural population activity (Bullock et al., 1995; Towle et al., 1999; Zaveri et al., 1999; Bruns and Eckhorn, 2004; Kristeva et al., 2007; DeCoteau et al., 2007a,b; Montgomery and Buzsáki, 2007; Sirota et al., 2008; Bollimunta et al., 2008), and the relationship between neuron spiking and field potentials (spike-field coherence) (Fries et al., 2001, 2008; Womelsdorf et al., 2006; Witham et al., 2007; Pesaran et al., 2008; Gregoriou et al., 2009; Jutras et al., 2009; Chalk et al., 2010). The spike-field coherence, $C_{ny}(f)$, can be defined in a fashion analogous to the definition of the "field–field", or more standard coherence between random processes modelling field-type recordings. This definition, discussed in more detail in Lepage et al. (2011), is summarized in the following. Let dn_t be the number of spiking events that occur in the time interval, $[(t-1)\Delta, t\Delta)$. Here Δ is the time between field measurements, and t is the time-index associated with the tth bin. The collection of these counts is a time-series and can be usefully modelled as a truncated realization[1] of a discrete-time point process, **dn**. Here, dn_t, is the tth element of **dn**, and is a random variable whose realization is the number of spiking events that occur in the interval $[(t-1)\Delta, t\Delta)$. To avoid multiple events in an observation interval, Δ is chosen sufficiently small such that the probability of multiple spiking events in any one observation interval is negligibly small. Note that this is possible for single neuron recordings due to the refractory period immediately following a neuron spiking event. During this refractory period, subsequent neuron spiking is greatly suppressed (Koch, 1999). A point process is completely characterized by its conditional intensity, Λ_t,

$$\Lambda_t = \lim_{\Delta \to 0} \frac{P(dn_t = 1 | H_t)}{\Delta}, \tag{1}$$

where H_t is the spike history process (Daley and Vere-Jones, 2003). Intuitively, the probability of an event at time t equals $\Delta \cdot \Lambda_t$, up to negligible corrections due to the small non-zero probability of multiple events in any one increment (Daley and Vere-Jones, 2003). When the increments, dn_t, do not depend on either past or future increments, the point process is called Poisson, and the conditional intensity, Λ_t, is equal to the rate of occurrence of spiking events. While sometimes convenient, this model is physiologically inaccurate due to dependence on past spiking. As described in Lepage et al. (2011), a spike-field coherence consistent with the more standard field–field coherence is defined in terms of weak-sense stationary random processes. Thus, the first two moments of dn_t must be independent of absolute time. Time-dependent firing activity, while maintaining stationary first and second moments can be attained by generalizing the discrete-time point-process to a doubly-stochastic discrete-time point process. That is, let the conditional intensity, Λ_t, be itself a weak-sense stationary random process such that $\Lambda_t \geq 0$. With this stipulation, both the point process modeling the spikes and the intensity, which determines the probability of a spike in each time-step, are both random processes. Let the centered increments of the discrete-time point process, dn_t be $d\tilde{n}_t$ such that,

$$d\tilde{n}_t = dn_t - E\{dn_t\}, \tag{2}$$

where E denotes the expectation operator. This ensures that $E\{d\tilde{n}_t\} = 0$. In analogy with the standard discrete-time Fourier transform, define the discrete-time Fourier transform of the centered increments, $d\tilde{n}_t$, evaluated at frequency f, as,

$$N_T(f) = \frac{\Delta}{T} \sum_{t=0}^{N-1} e^{-i2\pi f t \Delta} d\tilde{n}_t. \tag{3}$$

Here T is the duration of the time-series and Δ is the duration between samples. Let the local-field potential recording be represented by the weak-sense stationary random process y_t, with an associated discrete-time Fourier transform, $Y_T(f)$,

$$Y_T(f) = \frac{\Delta}{T} \sum_{t=0}^{N-1} e^{-i2\pi f t \Delta} y_t. \tag{4}$$

The spike-field coherence between the spiking and the local-field potential, $C_{\tilde{n}y}(f)$, is

$$C_{\tilde{n}y}(f) = \lim_{T \to \infty} \frac{E[N_T(f) Y_T^*(f)]}{\sqrt{E[|N_T(f)|^2] E[|Y_T(f)|^2]}}. \tag{5}$$

If the relevant spectra exist, Eq. (5) can be re-written,

$$C_{\tilde{n}y}(f) = \frac{S_{\tilde{n}y}(f)}{\sqrt{S_{\tilde{n}\tilde{n}}(f) S_{yy}(f)}} \tag{6}$$

where $S_{\tilde{n}y}(f)$ is the cross-spectrum between $d\tilde{n}_t$ and y_t, $S_{\tilde{n}\tilde{n}}(f)$ is the spectrum of $d\tilde{n}_t$ and $S_{yy}(f)$ is the spectrum of y_t. Through the Cauchy–Schwartz inequality, $0 \leq |C_{\tilde{n}y}(f)| \leq 1$ and $|C_{\tilde{n}y}(f)| = 1$ when there is a linear relation between $N_T(f)$ and $Y_T(f)$.

[1] Random processes which begin and end are not weak-sense stationary. That is, the dependence changes at the beginning and end of the random process and hence there is dependence on absolute time. Actual recordings begin and end, and are typically handled by realizing an infinite time-series and then truncating this realization to the recording duration.

Given the spike-field coherence defined in Eq. (6), the following relation between the spike-field coherence, $C_{ny}(f)$ and the intensity-field coherence $C_{\Lambda y}(f)$ exists,[2]

$$C_{ny}(f) = \frac{C_{\Lambda y}(f)}{\sqrt{1 + \frac{\mu_\Lambda + H(f)}{S_{\Lambda\Lambda}(f)}}} \tag{7}$$

where $\mu_\Lambda = E\{dn_t\}$ is the expected conditional intensity or rate of the spike process, $C_{\Lambda y}(f)$ is the coherence between the conditional intensity and the field potential, $S_{\Lambda\Lambda}(f)$ is the spectrum of the rate Λ_t, and $H(f)$ is a parameter influenced by history dependent spiking (such as a refractory period or periods of bursting). The behavior of this parameter is discussed in Section 5.4 of Lepage et al. (2011) for different types of neural activity, and is shown to be zero for spiking activity without history dependence. Thus, the spike-field coherence, $C_{ny}(f)$, depends on the two mean-square continuous processes Λ_t and y_t, and the mean rate of neural firing μ_Λ.

The dependence of the spike-field coherence $C_{ny}(f)$, on the expected number of spikes per bin, μ_Λ, renders the spike-field coherence responsive not only to the degree of association between the field rhythm and the spike times; but, also to the overall neuron activity. When attempting to compare the degree of associativity between two pairs of spike trains and local field potential rhythms, this confound, present when the total number of spikes differs between the pairs, has been dealt with in the literature by, for example, randomly thinning the spike times of the more active neuron until the neurons possess the same number of spikes in both pairs (Mitchell et al., 2009). While not the focus of this paper, the procedure is included in the simulation study presented in Section 4 for comparison with the proposed techniques. In Section 3 a method of comparing spike-field coherence across spike-time local-field potential pairs is introduced. This method, capable of disambiguating changes in spiking rate from changes in spike-field association, is demonstrated in simulation in Section 4 and on real data in Section 5.

2.1. Rhythmic spike-field parametric modelling

In this paper, the confound confusing LFP-rhythm-spike-time-associativity with overall neuron activity is addressed with the aid of a test between parameters in a class of parametric models. This procedure is described in Section 3. In the class of generalized linear models, nonlinear link functions relate linear combinations of covariates to a neuron's expected intensity or rate (Truccolo et al., 2005). In this work, along with the log-link function employed in Truccolo et al. (2005) to relate covariates to spiking activity, a piecewise linear link function is also used. These two link functions, both necessarily nonlinear due to the non-negativity of neural spiking rates, mix LFP rhythms at frequencies other than the frequency of interest with the frequency of interest in different ways. The piece-wise linear link function is chosen such that the theoretical advantages of the Fourier basis can be exploited in the current context. In particular, as shown in Appendix C, model mis-specification due to unmodelled sinusoids with different frequencies has limited effect when firing rates are sufficiently high and a sufficiently narrow bandwidth is chosen when estimating LFP phase. The log-link function, while more susceptible to effects due to unmodeled sinusoids, is capable of describing broad-band firing phenomenon such as greater preferred-phase spike time tuning and is essentially a von-Mises distribution on the LFP phase associated with the time of a neural spike.

The proposed test, described in Section 3, is based upon the standard LFP cosine-tuning parametric model of neural spiking, used for example in studies of the motor cortex (Sanger, 1994). This cosine-tuning model relates a single sinusoid plus background rate to the expected stochastic conditional intensity. As before, let the stochastic conditional intensity at time-index t be Λ_t. Model the stochastic conditional intensity of the doubly-stochastic discrete time point process as,

$$\Lambda_t = g^{(j)}(\tilde{\Lambda}_t), \tag{8}$$

where

$$g^{(j)}(x) = \begin{cases} \max(0, x), & j = \text{piecewise linear link} \\ e^x, & j = \text{log link} \end{cases}$$

and

$$\tilde{\Lambda}_t = \alpha + \beta_c \cos(\phi_t) + \beta_s \sin(\phi_t) + \sum_{k=1}^{K} \gamma_k dn_{t-k}. \tag{9}$$

Here ϕ_t is the instantaneous phase of LFP rhythm for a frequency of choice, α is the theoretical background, and the effect upon the instantaneous stochastic conditional intensity due to past spiking is incorporated in the linear combination of past spiking intervals: $\sum_{k=1}^{K} \gamma_k dn_{t-k}$.[3] In this paper, background rate and the theoretical background are synonymous; while average rate refers to an estimate of the theoretical background rate computed as the sum of spiking events divided by the observed duration. When all of the γ_k variables are zero, the stochastic conditional intensity has a degenerate probability distribution such that Λ_t is deterministic, and further, because there is no dependence on past (or future) spiking activity, the point process becomes Poisson in the sense that the increments, dn_t, are independent of each other. The modulation, ρ, of the stochastic conditional intensity due to the LFP rhythm is given by,

$$\rho = \sqrt{\beta_c^2 + \beta_s^2}. \tag{10}$$

The preferred phase of LFP rhythm, ϕ_p, at which neuron spikes occur is,

$$\phi_p = \tan^{-1}\left(\frac{\beta_s}{\beta_c}\right). \tag{11}$$

As in Truccolo et al. (2005), estimation is performed by maximizing the likelihood. When $g(\tilde{\lambda})$ in Eq. (8) is the log link, then the log link is canonical, and more standard GLM model fitting routines can be applied. When $g(\tilde{\lambda})$ is piecewise linear, however, a modified Newton–Raphson algorithm is used to maximize the likelihood. A description of the Newton–Raphson procedure for the model described by Eqs. (8) and (9) is given in Appendix A. Subsequent inference can be conducted on the parameter estimates themselves by employing the asymptotic distribution of the parameter estimators, $\hat{\beta}_j$ (Truccolo et al., 2005), or through model comparison with information measures such as the Akaike information criterion

[2] The auto-covariance sequences and the cross-covariance sequences are identical when computed with either the centered increments, $d\tilde{n}_t$, or the non-centered increments, dn_t. The centering occurs in the definition of these sequences. Under standard conditions, the auto-spectra and cross-spectrum are discrete-Fourier transforms of these sequences. Then $C_{\tilde{n}y}(f)$ is equal to $C_{ny}(f)$.

[3] A doubly-stochastic point process is one where the intensity, and the times at which spiking events occur are both random processes, while discrete-time refers to the duration over which events are counted and a single value reported representing the sum of spiking events that occurred during the bin interval. A realization of such a point process proceeds bin-wise from early times to late times. First the intensity is realized, after which spike times are determined according to a Poisson probability mass function with a rate equal to the realized intensity. Any history effects due to spikes in the first bin affect the realized intensity at the second bin and the process continues until spike counts for all bins are determined.

(Akaike, 1974), or by comparing models via the asymptotic distribution of the likelihood ratio statistic computed between nested models (Casella and Berger, 2001). For a recent application of these latter two techniques see MacDonald et al. (2011).

It is important to note that Eqs. (8) and (9), when using the piecewise linear link function, completely captures, under assumptions defined in Appendix C, the different ways that a narrow-band LFP rhythm can influence spiking activity. That is, under a change in condition, either the modulation, ρ, can change, and/or the preferred phase of spiking, ϕ_p, can change. However, the degree to which neural spiking activity is locked to a specific phase of LFP rhythm is not a narrow-band phenomenon. This neural behaviour, when firing is "tuned" to a specific phase of LFP rhythm is addressed by using the log link function. In this latter case, information across frequency harmonics is combined to assess the extent to which neural firing activity occurs in a temporally "punctate" fashion, periodically, at a specific phase of LFP rhythm. This phenomenon is further discussed in Section 3.1.

With the model specified by Eqs. (8) and (9), an analogous measure similar to spike field coherence is available by estimating the β_c and β_s parameters; large values of ρ indicate a strong relationship between LFP rhythm and neural spike times. To determine the significance of this modulation, the β_s and β_c parameters can be tested for significance by using the asymptotic result that the associated parameter estimators, $\hat{\beta}_s$ and $\hat{\beta}_c$ are normally distributed and possess a link-dependent covariance matrix, $\mathbf{C}_\beta^{(j)}$, asymptotically equal to the negative of the inverse of the observed Fisher information, $\mathbf{I}_\beta^{(j)}$,

$$\mathbf{C}_\beta^{(j)} = -(\mathbf{I}_\beta^{(j)})^{-1} \tag{12}$$

$$\mathbf{C}_\beta^{(j)} = \left[\mathbf{H}^T\mathbf{D}^{(j)}\mathbf{H}\right]^{-1}. \tag{13}$$

Here $\mathbf{H}$ is a model matrix, with a kth row equal to, $(\mathbf{H})_{tj} = [1\ \cos(\phi_t)\ \sin(\phi_t)]$, and

$$\mathbf{D} = \begin{cases} \operatorname{diag}\left\{\dfrac{\mathbf{n}}{\check{\boldsymbol{\lambda}}^2}\right\}, & j = \text{piecewise linear link} \\ \operatorname{diag}\left\{\check{\boldsymbol{\lambda}}\right\}, & j = \text{log link} \end{cases} \tag{14}$$

is a diagonal matrix. Here, the kth element of $\mathbf{n}$ is equal to n_k the number of observed counts within bin k. The maximum likelihood estimator of the stochastic conditional intensity, $\boldsymbol{\Lambda}$, is $\hat{\boldsymbol{\lambda}}$, and $\check{\boldsymbol{\lambda}}$ is the maximum likelihood estimate; equal to a realization of the random maximum likelihood estimator, $\hat{\boldsymbol{\lambda}}$. The division within the *diag*( ) operator is element-by-element division. While the asymptotic convergence to the normal distribution is established when the link function is twice differentiable (McCullagh and Nelder, 1999), the piecewise linear link is not differentiable at the origin. In Appendix B, it is established that the asymptotic approximation is acceptable for standard sized neuroscience data sets with the non-differentiable piecewise linear link specified in Eq. (9).

In this estimation scheme, the instantaneous phase of the LFP rhythm for some frequency of interest, f_0, is considered known. In fact, the instantaneous phase of the LFP rhythm is not known and must be estimated from the local field potential time-series. In this paper, this estimate is computed using the analytic signal of the bandpass filtered local field potential in a procedure often called the "Hilbert transform" phase estimate (Bruns, 2004). Maximum likelihood estimation assuming knowledge of the LFP instantaneous phase is maximum conditional likelihood estimation, where the extremization is of the likelihood conditioned on the instantaneous LFP phase equalling the estimated instantaneous phase; rather than an extremization of the full likelihood. Thus, the bias and variance properties of the conditional maximum likelihood estimators may differ from those of the maximum likelihood estimator. While a concern, with a sufficient number of trials and with an appropriately chosen bandpass filter when estimating the instantaneous field phase[4] one expects the maximum conditional likelihood estimator to be similar to the maximum likelihood estimator. In the large data limit, successful inference can be conducted with the phase estimate conditioned likelihood.

2.2. Model mis-specification

A common issue regarding parametric modelling is the problem of model selection. While model selection techniques are discussed in Truccolo et al. (2005) and careful analysis includes their use; it is worth noting that model mis-specification is common (Box and Draper, 1987), and that inference with mis-specified models is often useful. When the data model, expressed by Eqs. (8) and (9), explains all features of the data there is no model mis-specification, and the resulting conditional maximum likelihood estimator, $\hat{\boldsymbol{\lambda}}$, discussed in Section 2.1, is approximately unbiased for a data size typical of neuroscience experiments. In this situation, $\hat{C}_\beta$ is unbiased. Further $\hat{C}_\beta$ and $\hat{\boldsymbol{\lambda}}$ are consistent; as they equal their respective theoretical quantities in the large data limit. When the data model fails to capture data features, model mis-specification occurs and may bias parameter estimators. Five potential sources of bias are (1) unmodelled history dependent neural spiking activity, (2) unmodelled neural spiking activity related to covariates such as experimental condition, (3) bias due to a mis-specified relationship between the covariates and the conditional intensity, (4) the effect of unmodelled sinusoids oscillating with frequencies different than the modelled frequency interval and (5) the effect of unspecified self-sinusoidal terms present in the conditional intensity. Reduction of the first form of bias, unmodelled effects of previous neural activity, can be accomplished by incorporating covariates representing past neural activity, as described in Eq. (9). Similarly, reduction of the second form of bias, effects due to unmodelled experimental covariates, can be accomplished by explicitly incorporating these covariates into Eq. (9). The third form of bias is more complicated and depends on the bandwidth used to estimate the instantaneous phase of the field-type time series, on the overall rate of spiking, and the functional form of the nonlinearity relating covariates to the conditional intensity. A detailed theoretical account of these latter effects is beyond the scope of this work. The fourth form of mis-specification is addressed in Appendix C for the situation where the link function is accurately specified to be the piecewise linear link function. It is shown that there is no bias incurred due to unmodelled sinusoids when the actual rate is always greater than zero when the LFP rhythm is sinusoidal. When neural firing rates are low, the nonlinear link can couple frequency components and introduce bias.[5] Low here means that oscillatory activity dips sufficiently below zero such that the background plus this oscillatory activity would be less than zero if not for the "capping" effect of the piecewise nonlinearity. Thus, for sparse neural activity, prominent spectral peaks should be explicitly incorporated into Eq. (9) as further sinusoid terms to account for their influence. When using the log link to relate the stochastic conditional intensity to neural activity this source of bias may exist for any rate of neural spiking. In this paper, discussion is restricted to the case where spiking rates are sufficiently high such that the

[4] That is, a band-pass filter with a passband specified to be the interval of frequencies required to describe the field rhythm. This specification is required to avoid phase estimator bias. Note that bias results from model mis-specification.

[5] Both link functions are nonlinear; and further, any link function that insures that the intensity is non-negative is nonlinear.

nonlinear portion of the piecewise linear link function is not a concern. While important, discussion of lower spiking rate regimes is left for future work. The final form of mis-specification is separable from the influence of the LFP rhythmic covariate in the situation where multiple trials are recorded. In this situation, the influence of oscillatory components in the conditional intensity occur at a phase random relative to the LFP phase, and are not expected to bias the mis-specified model, Eqs. (8) and (9). While one does not expect bias, note that the relative effect of LFP rhythmic influence to self-oscillation on spiking activity can be quantified with a correctly specified model formed by adding oscillatory terms to Eqs. (8) and (9) that are not linked to the phase of the LFP rhythm, and possess a trial dependent phase offset. This model allows, for example, the per-frequency study of the importance of LFP influence on spiking activity relative to the influence of self oscillation on spiking activity.

One notes that inference with the model specified by Eqs. (8) and (9) typically involves the covariance matrix estimate specified in Eq. (13). In Eq. (13), the diagonal matrix, $\mathbf{D}$, involves the unknown, theoretical conditional intensity, $\boldsymbol{\lambda}$, along the diagonal. In practice, an estimate of $\mathbf{C}_{\beta}^{(j)}$ is computed using the estimated conditional intensity or rate. Due to the dependence of the covariance matrix estimator, $\hat{\mathbf{C}}_{\beta}^{(j)}$ upon the rate estimator, $\hat{\boldsymbol{\lambda}}$, the covariance matrix estimator can also be biased. While this possibility exists, in Appendix C it is shown when employing the piecewise linear link that the estimate of the covariance matrix employed in this work is not affected by unmodelled sinusoids influencing neural activity for neural firing rates that are strictly greater than zero.

2.3. Frequency interval specification

Prior to performing the test presented in Section 3, it is necessary to specify the frequency interval at which to test for changes in coupling between spike times and LFP rhythm. This interval is specified either through a priori information or through a preliminary analysis designed to find frequency intervals of interest. Subject to computational resources and desired statistical power, one notes that the test for condition preferential spike-time LFP rhythmic coupling can be applied sequentially to frequency intervals spanning zero frequency to Nyquist frequency. An example of this procedure will be described later in Section 5 and Fig. 9. In this latter scenario it is advisable to employ appropriate multiple hypothesis testing techniques.

3. Test for equality of spike-field association

To test for a difference in spike-field association between two pairs of LFP rhythm/neural spike times, the model specified by Eqs. (8) and (9) is fit to each of the neural spike time/LFP rhythm pairs. The resulting parameter estimators,

$$\hat{\beta}_c^{(k)}, \hat{\beta}_s^{(k)}, \quad k = 1, 2, \tag{15}$$

are the amplitude of the cosine of the instantaneous LFP phase, and the amplitude of the sine of the instantaneous LFP phase. The integer, k, specifies which of the two sets of pairs to associate with the estimators. The distribution of these estimators asymptotically converges to a normal distribution with a mean equal to the theoretical mean, and a variance given by the covariance matrix estimator, $\hat{C}_{\beta}^{(k)}$. This asymptotic convergence is further discussed in Appendix B. Associated with these estimators are estimators for the modulation.

$$\hat{\rho}^{(k)}, \quad k = 1, 2. \tag{16}$$

Under the null hypothesis that the modulation, ρ, of the spiking activity with the LFP rhythm is the same between $k = 1$ and $k = 2$, the difference of the theoretical modulations is zero,

$$\begin{aligned} d_\rho &= \rho^{(1)} - \rho^{(2)}, \\ &= 0, \end{aligned} \tag{17}$$

with an associated estimator, $\hat{d}_\rho$,

$$\hat{d}_\rho = \hat{\rho}^{(1)} - \hat{\rho}^{(2)}. \tag{18}$$

For the kth condition, $\hat{\rho}^{(k)}$ has an asymptotic Rice, probability density function, $f_{\hat{\rho}^{(k)}}(x)$ (Rice, 1945),

$$f_{\hat{\rho}^{(k)}}(x) = \frac{x}{\sigma_k^2} I_0\left(\frac{x\rho^{(k)}}{\sigma_k^2}\right) e^{-\frac{1}{2\sigma_k^2}(x^2+(\rho^{(k)})^2)}. \tag{19}$$

Here I_0 is the zeroth order modified Bessel function of the first kind, σ_k^2 is the variance of $\hat{\beta}_c^{(k)}$ and of $\hat{\beta}_s^{(k)}$ and equals the appropriate diagonal element of the observed parameter covariance matrix, $\check{C}_{\beta}^{(k)}$. Here the $\check{}$ symbol denotes the observed quantity computed from data. Note that with the piecewise linear link function, in the narrowband, sufficiently high rate regime discussed in Section 3.1, the variance of $\beta_c^{(k)}$ and the variance of $\beta_s^{(k)}$ are identical. In practice, σ_k^2 is taken to be the average of the appropriate diagonal elements of $\check{C}_{\beta}^{(k)}$. When using the log link, these variances are found empirically to be similar, and the same average is employed.

Under the null hypothesis of no change in modulation, the modulation difference, d_ρ, is zero,

$$d_\rho = \rho^{(1)} - \rho^{(2)}, \tag{20}$$

and the difference estimator, $\hat{d}_\rho$, is a random variable equal to the sum of a Rice distributed random variable with mean, $E\{\hat{\rho}^{(1)}\}$, and variance σ_1^2 with a second independent random variable. This second random variable is the negative of a Rice distributed random variable with mean equal to $E\{\hat{\rho}^{(2)}\}$ and variance σ_2^2. This resulting distribution under the null hypothesis is computed numerically by convolving the relevant distributions and the reported two-sided test p-value, p_{value}, is,[6]

$$p_{\text{value}} = P(\hat{d}_\rho > |\check{\rho}^{(1)} - \check{\rho}^{(2)}|) + P(\hat{d}_\rho < -|\check{\rho}^{(1)} - \check{\rho}^{(2)}|), \tag{21}$$

where $\check{\rho}^{(k)}$ is the observed modulation and is a realization of the estimator, $\hat{\rho}^{(k)}$, for $k = 1, 2$.

When p_{value} is very small, numerical issues can arise due to the necessity of accurately quantizing a large range of possible values of $\hat{d}_\rho$. To avoid excess computational burden, numerical difficulties are detected by checking the Riemann approximation to the integral of the probability density function for $\hat{d}_\rho$ under the null hypothesis of no across-condition change. When this integral deviates substantially from 1, to accurately represent the probability density functions numerically, a fine grid of points spanning a large interval is required. Thus, the occurrence of numerical difficulty is restricted to the case where $p_{\text{value}} \ll 0.01$, and a more conservative and computationally efficient procedure is used. In essence, this latter testing procedure is a principled way of assigning a conservative yet small p-value for strong detections that would otherwise require a prohibitively expensive computation. This second test, based upon an inequality due to Cantelli (Feller, 1966), computes the p_{value} as follows,

$$p_{\text{value}} = \max\{UB_1, UB_2\}, \tag{22}$$

[6] The sum of two independent random variables is a third random variable with a probability density function (when it exists) equal to the convolution of the probability density functions of the summed random variables (Casella and Berger, 2001).

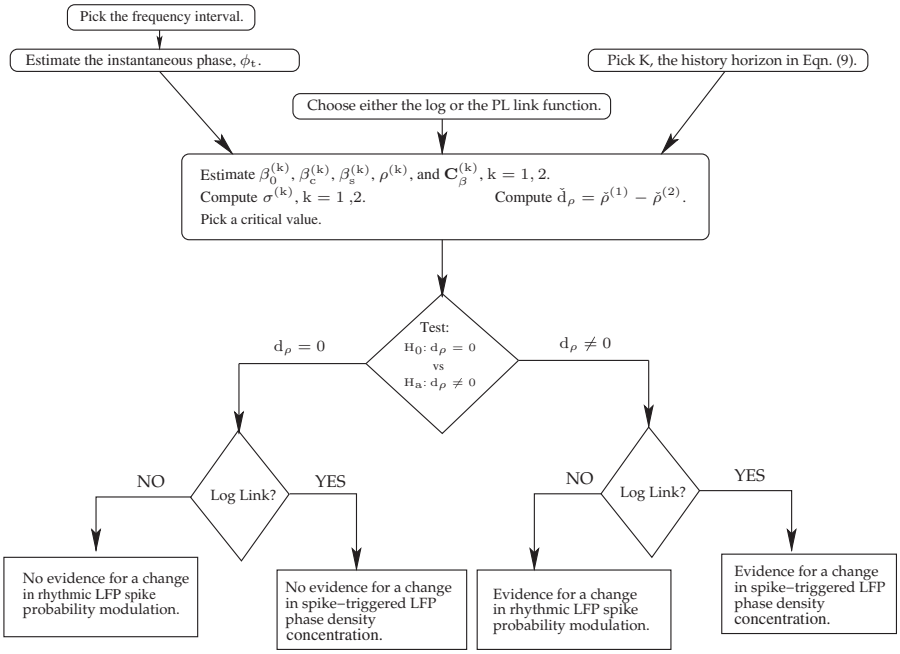

Fig. 1. Testing procedure and associated inference. The flow chart describes the key steps necessary to perform the proposed change-in-modulation testing procedure for a single frequency interval.

where the upper bound, UB_1, on the probability of exceeding a modulation difference of $\breve{d}_\rho$, given a modulation standard deviation equal to σ_1, is,

$$UB_1 = \frac{1}{1+(\breve{d}_\rho/\sigma_1)^2}, \tag{23}$$

and the upper bound, UB_2, on the probability of exceeding a modulation difference of $\breve{d}_\rho$, given a modulation standard deviation equal to σ_2, is,

$$UB_2 = \frac{1}{1+(\breve{d}_\rho/\sigma_2)^2}. \tag{24}$$

This testing procedure, involving these two tests employed adaptively depending on the data, is applied to synthetic data in Section 4 and to real data in Section 5, and breaks the associativity-spiking-rate ambiguity present in spike-field coherence. See Fig. 1 for a schematic description of the steps required to implement the proposed testing procedure.

In some cases it may be desirable to test for differences in the parameter α. As in Eq. (15), the model is fit to the data in both conditions, and the α estimators are indexed by the condition number, k,

$$\alpha^{(k)}, \quad k = 1, 2. \tag{25}$$

Each of these quantities is normally distributed with a variance given by the first diagonal element of $\breve{\mathbf{C}}_\beta^{(k)}$. Let this variance be $\sigma^2_{\alpha^{(k)}}$. Under the null hypothesis that there is no change in the background spiking across condition, the difference, $\hat{d}_\alpha$, is normally distributed with a mean of zero and a variance, $\sigma^2_{\hat{d}_\alpha}$ equal to the sum of the variances of $\hat{\alpha}^{(k)}$,

$$\hat{d}_\alpha \sim N(0, \sigma^2_{\alpha^{(1)}} + \sigma^2_{\alpha^{(2)}}). \tag{26}$$

For a specific observed difference, $\breve{d}_\alpha$, equal to the difference, $\breve{\alpha}^{(1)} - \breve{\alpha}^{(2)}$, the two-sided test p-value, p_{value}, is,

$$p_{\text{value}} = P(\hat{d}_\alpha > |\breve{d}_\alpha|) + P(\hat{d}_\alpha < -|\breve{d}_\alpha|). \tag{27}$$

3.1. Piecewise linear and log link functions: testing differences

The behaviour of the test computed using the log-link is determined by the effect of a change in modulation when using the log-link. To illustrate this fact, Eq. (8) can be re-written, ignoring history effects, to be proportional to a von-Mises (Mardia and Jupp, 2000) probability density[7]

$$\begin{aligned}\Lambda_t &= e^{\alpha+\beta_c\cos(\phi_t)+\beta_s\sin(\phi_t)},\\ &= c\, e^{\rho\cos(\phi_t-\phi_p)},\end{aligned} \tag{28}$$

where $c = e^\alpha$ and ρ and ϕ_p are defined in Eqs. (10) and (11), respectively. Thus, when ρ in Eq. (28) is large, the rate, Λ_t is sharply attenuated as ϕ_t deviates from ϕ_p, and similarly, when ρ equals zero, LFP rhythm phase has no effect on spiking probability, and the phase distribution is uniform. Here, if the background, α is large, then c is large, and the spiking rate increases but the shape of the LFP phases at spike times remains un-altered. This situation is in stark contrast to the rate obtained with the piecewise linear model. For $\alpha > \rho$, the piecewise linear link is in a linear regime, and the LFP phase at spike times is sinusoidally distributed about the period with a peak at ϕ_p. As α increases the LFP rhythm phases at the time of spiking tend to occur randomly during the LFP rhythm period and the phase density tends to a uniform probability density. These differences are exemplified by the results of the simulation depicted in Fig. 4 in Section 4. The testing procedure, and the differing interpretations associated with the piecewise-linear (PL) and log link functions, are shown in a flow chart presented in Fig. 1. There are two tests and four possible outcomes. Table 1 provides qualtitative interpretation for these four outcomes.

3.2. Relation to spike-field coherence

The proposed test computed with the piecewise linear (PL) link function is similar to an across-condition comparison of rhythmic

[7] This is a unimodal function of ϕ_t centered upon the preferred phase, ϕ_p, with a width specified by ρ.

Table 1
Interpretation of testing results. Let $d_\rho^{(log)}$ and $d_\rho^{(PL)}$ denote, respectively, the across-condition change in modulation associated with the log link, and the PL link. The four test outcomes are described below for a single frequency interval.

Test results	Interpretation
$d_\rho^{(PL)} = 0, d_\rho^{(log)} = 0$	No evidence for a condition dependent rhythmic LFP influence.
$d_\rho^{(PL)} = 0, d_\rho^{(log)} \neq 0$	Evidence for a condition dependent spike-triggered phase density concentration that does not affect rhythmic LFP influence for this frequency interval.
$d_\rho^{(PL)} \neq 0, d_\rho^{(log)} = 0$	Evidence for a condition dependent rhythmic LFP influence that arises by drawing a condition dependent number of spikes from the same spike-triggered phase density.
$d_\rho^{(PL)} \neq 0, d_\rho^{(log)} \neq 0$	Evidence for a condition dependent rhythmic LFP influence and a condition dependent spike-triggered phase density concentration.

spike-field association based upon spike-field coherence, Eq. (6). This similarity arises because, on a per frequency interval basis, for high rates, both of these tests will respond to the strength of coupling between the LFP rhythm and the conditional intensity or rate. This is not the situation with the log link function, which responds to the strength of spike activity tuning to a preferred LFP rhythm phase. One notes however, that if the data are generated according to the PL model, changes in the ratio of α to ρ will affect spike tuning to the preferred phase; though tight tuning to a specific phase is not possible. If the data are generated with the log link, changes in the α parameter will affect the inferred modulation, ρ, if inference is performed with the piecewise linear link. Both spiking models are capable of generating data that is uncoupled from the LFP rhythm and has no preferred LFP rhythm phase of spiking. For sufficiently high rates, only the model using the log link can tightly tune spike times to a specific LFP rhythm phase (this

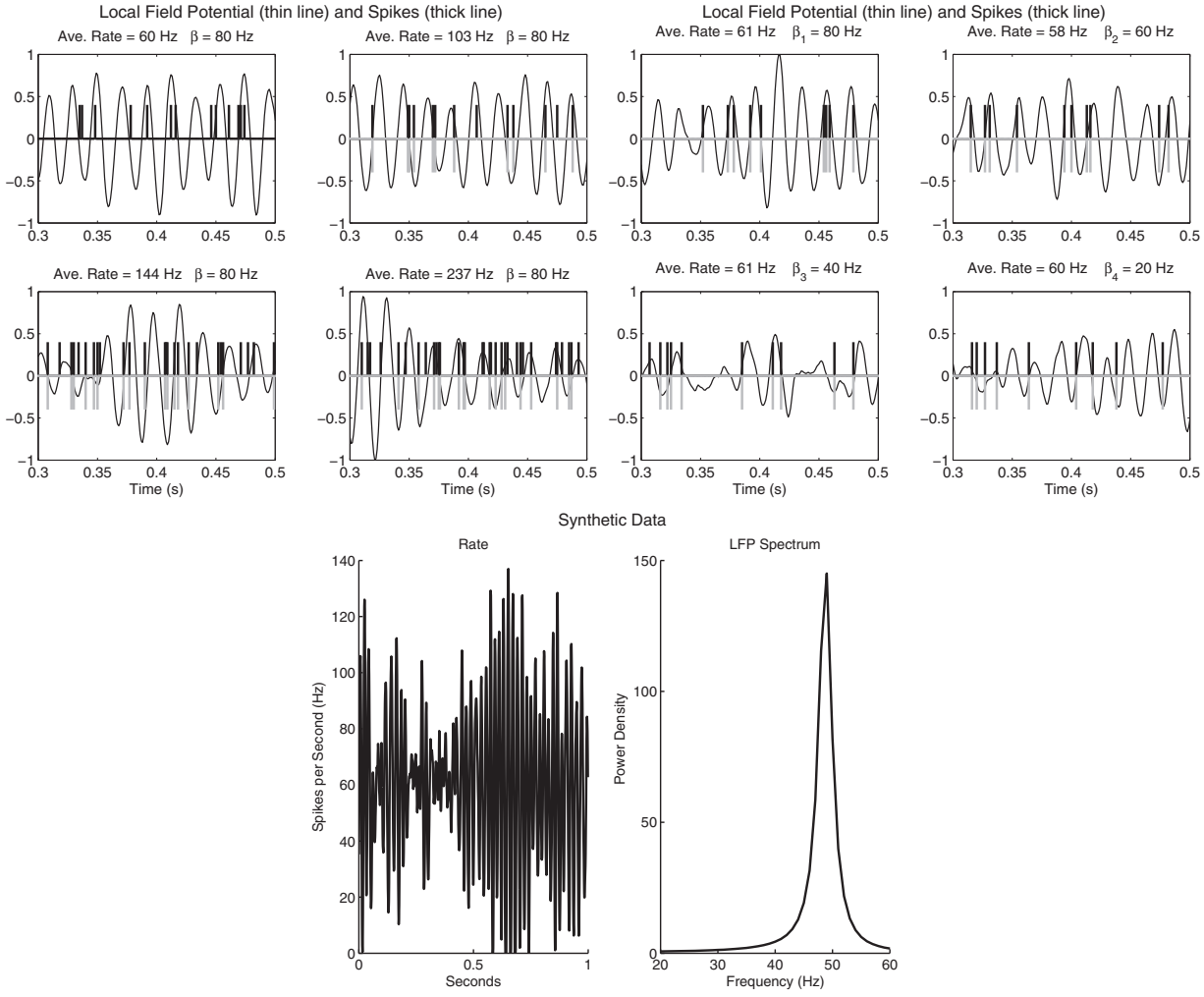

Fig. 2. Description of synthetic data used in Fig. (3). *Upper left half*: Example spike trains and LFP timeseries associated with the left hand side of Fig. (3). The link between the LFP and the spiking rate is constant, while the background rate, α_k, increases from left to right and from top to bottom. Rates are prevented from falling below zero by employing the piecewise linear link function. The preferred phase of spiking in all plots, ϕ_p, is zero. The dark thick lines indicate the spike times, and increase in occurrence frequency with increasing background rate. Note that the fraction of spike times that occur at the LFP peak decreases with increasing background rate. For $k > 1$, that is for all but the upper left plot, the thick non-positive gray line indicates the occurrence of spikes that have not been removed in the thinning procedure employed in Mitchell et al. (2009) and Gregoriou et al. (2009). Here spike trains are thinned pairwise between the spike trains corresponding to $k = 2, 3, 4$, and to the spike trains associated with $k = 1$. As expected, the times at which spikes occur at LFP peaks is greatly reduced in the thinned spike train depicted in the lower right hand plot for the $k = 4$ case. *Upper right half*: Example spike trains and LFP timeseries associated with the right hand side of Fig. (3). The link between the LFP and spiking rate decreases from left to right and top to bottom, while the background rate remains constant. *Bottom left*: One of twenty realizations of the spiking rate associated with the upper left plot in Fig. (2). *Bottom right*: The theoretical LFP spectrum associated with the synthetic rate plotted on left. The LFP is oscillatory with a spectral peak near 50 Hz. Realizations of this random process are used to generate all of the synthetic LFPs used in simulation.

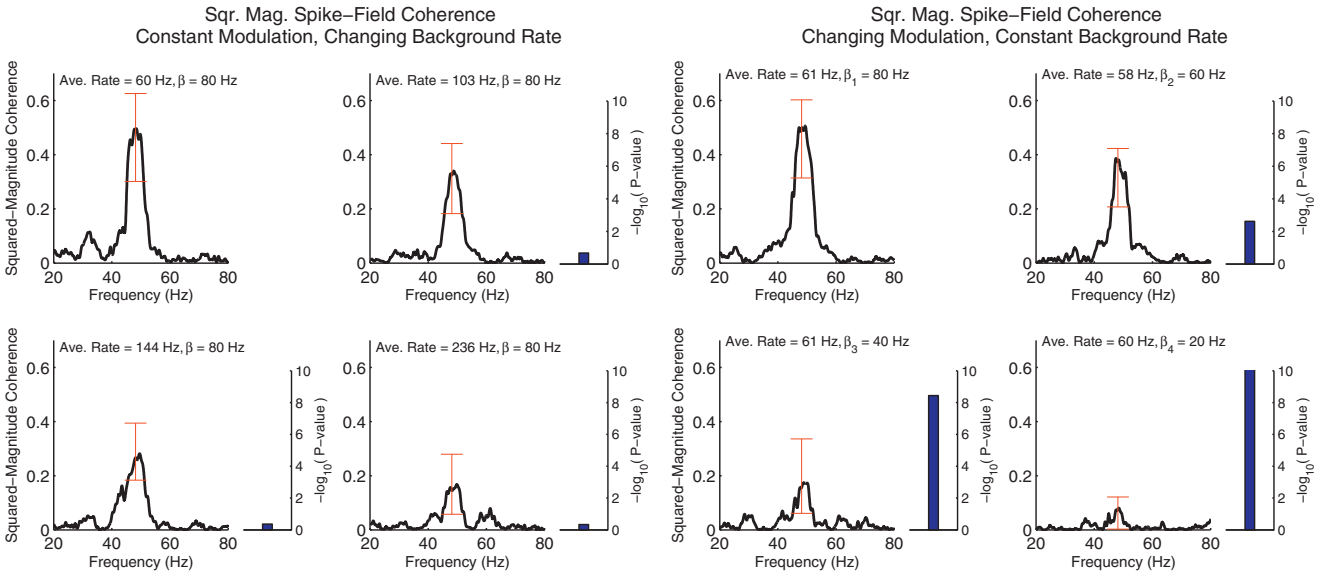

Fig. 3. Proposed test breaks average rate/association confound. The squared magnitude of the spike field coherence computed between spike trains and LFP time series for synthetic data generated using the model specified by Eq. (29) and illustrated by Fig. 2. When the spike probability modulation due to the LFP is held constant the squared magnitude of the spike field coherence decreases with increasing background rate (*left four plots*). When the background rate, α_k is held constant while the modulation, in this case equal to β_k, decreases, the squared magnitude of the spike field coherence decreases. Increasing background rate and decreasing association between neural spike times and LFP rhythm are indistinguishable in spike-field coherence. Bar plots: Negative of the base 10 logarithm of the *p*-values of the proposed test using the piecewise-linear link (PL), see Section 3, for differences in LFP rhythm/spike train association. The test is computed between the data used to compute the upper left plot and the data used to create the other three plots for each of the two cases (constant background rate (*left half*) and non-constant background rate (*right half*)). The test disambiguates changes in background rate from changes in the association between LFP rhythm and neural spike times. Ninety-nine percent confidence intervals are depicted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

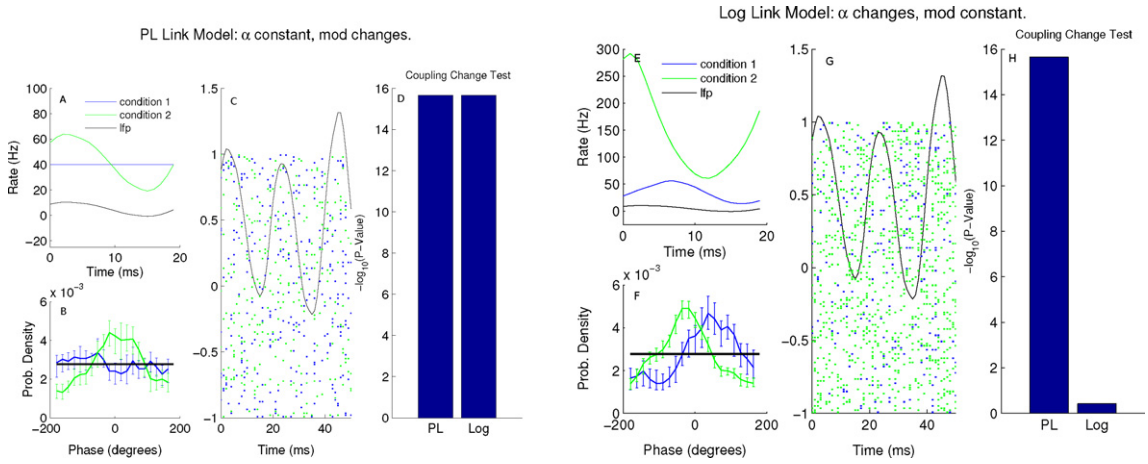

Fig. 4. Impact of PL and log link functions on tests of association. Both simulations: The LFP is as used in the first simulation and is depicted in Fig. 2. *Left-half (A–D)*: The modulation, ρ, is changed across conditions while the background α is held constant. The spike-train data are generated using the piecewise linear link (PL), example in plot *C*. The modulation changes across conditions and both tests are significant (*D*). *Right-half (E–H)*: The spike-train data are generated using the log link, the preferred phase of spiking changes across conditions, the spike-triggered LFP phase densities are otherwise equivalent (*F*), and the modulation is much stronger as demonstrated by the response of the rate as a function of LFP phase (*E*). Only the test computed using the piecewise linear link function is significant (*H*). This behaviour is consistent with the description in Section 3.1 of the difference between the test computed using the piecewise linear link function and the log link function. The ninety-five percent phase density confidence intervals are computed by boostrapping the trials. They have not been corrected for multiple comparisons. The black horizontal line is the uniform probability density function over phase.

is a broadband phenomenon[8]), and only the piecewise linear link effectively decouples LFP rhythmic coupling to the rate from background activity, since with the log link α controls the number of spikes drawn from the specified von-Mises type distribution and does not allow for a non-rhythmic rate offset. Note that this does not mean that inference regarding the modulation of spiking activity by LFP rhythm using the log link function will be sensitive to changes in the average spiking rate. These changes are still accommodated by the α parameter. However, in general, these changes are made by scaling an oscillatory rate rather than by shifting one.

Unlike spike-field coherence, the proposed tests separate, with different link-function dependent interpretations, the parameters linking LFP rhythm to spiking rate from the average rate of spiking.

[8] Here broadband means specifically the involvement of harmonics of the frequency located at the center of the narrowband frequency interval under consideration. These harmonics are required to produce a periodic rate that is not sinusoidal.

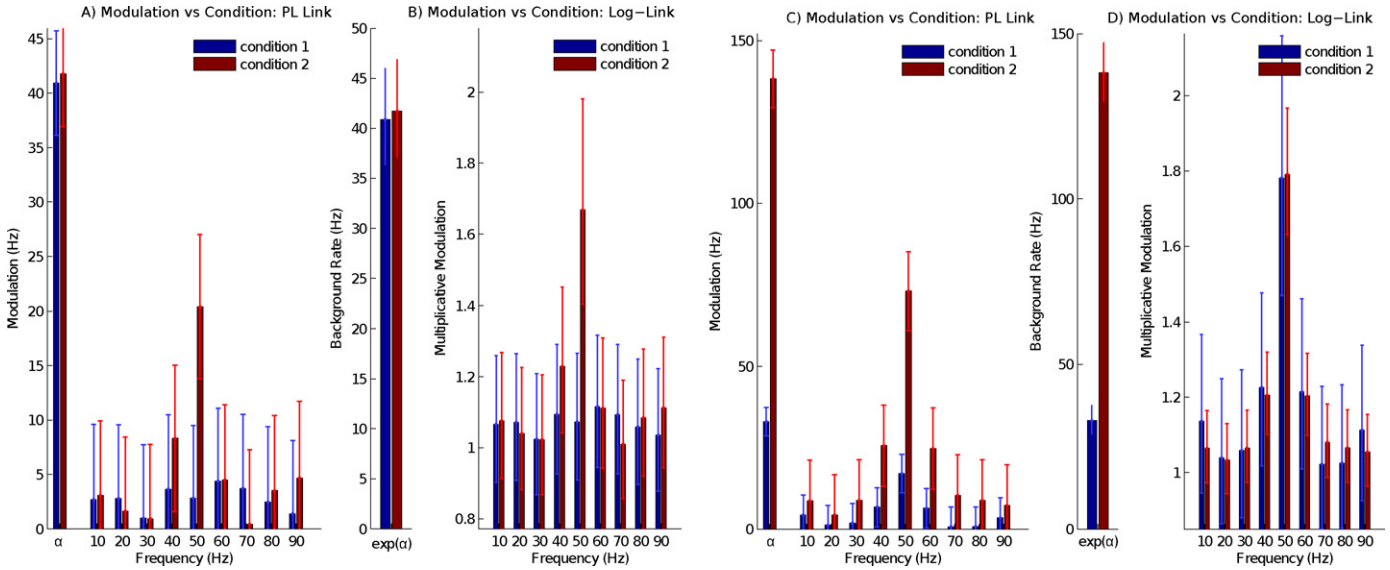

Fig. 5. Modulation, ρ, estimated from the data depicted in Fig. 4. All plots: LFP as described in Fig. 2. *Left-half (A,B)*: Per-frequency rate modulation, ρ, as well as background rate, α, estimated from the spiking data depicted in Fig. 4, plot C. This spiking data is generated with use of the PL link function and is used to produce the plots, (A–D) of Fig. 4. *Right-half (C,D)*: Per-frequency rate modulation, ρ, as well as background rate, α, estimated from the spiking data depicted in Fig. 4, plot G. This spiking data is generated with use of the log link function and is used to produce the plots, (E–H) of Fig. 4. Note that the interpretation of the ρ and α parameters depends upon the type of link function used. When the log link is employed, the modulation is multiplicative; e^{ρ}, multiplies a background rate of $\Delta^{-1}e^{\alpha}$ *Hz* to yield the modulation due to LFP rhythm in units of *Hz*. The data generated with the log link is significantly modulated at a number of frequencies about 50 Hz (C and D) for both conditions, but only modulation ρ associated with the PL statistical model relating LFP rhythm to spiking is significantly different across conditions. The confidence intervals are approximate, 95 percent confidence intervals computed with the delta method (see Appendix D). These intervals are Bonferroni corrected for multiple comparisons. The results are consistent with the results of the proposed hypothesis test plotted in Fig. 4.

4. Simulation

To illustrate the efficacy of the proposed tests two simulations are performed. In the first simulation, synthetic data for eight pairs of neuron spike times and LFP time-series are generated. The first four of these pairs link the neural spiking rate to an LFP via a constant scalar multiple, β, plus an LFP-independent background rate, α_k, that increases with spike train-LFP pair index, k. This index, k, indexes into the the set of four spike-train/LFP pairs, depicted in the left-hand-sides of Fig. 2. The spiking rate for the kth spike-train LFP pair, $\lambda_t^{(k)}$, is specified as

$$\begin{aligned}\lambda_t^{(k)} &= \max(0, \tilde{\lambda}_t^{(k)}),\\ \tilde{\lambda}_t^{(k)} &= \alpha_k + \beta y_t.\end{aligned} \tag{29}$$

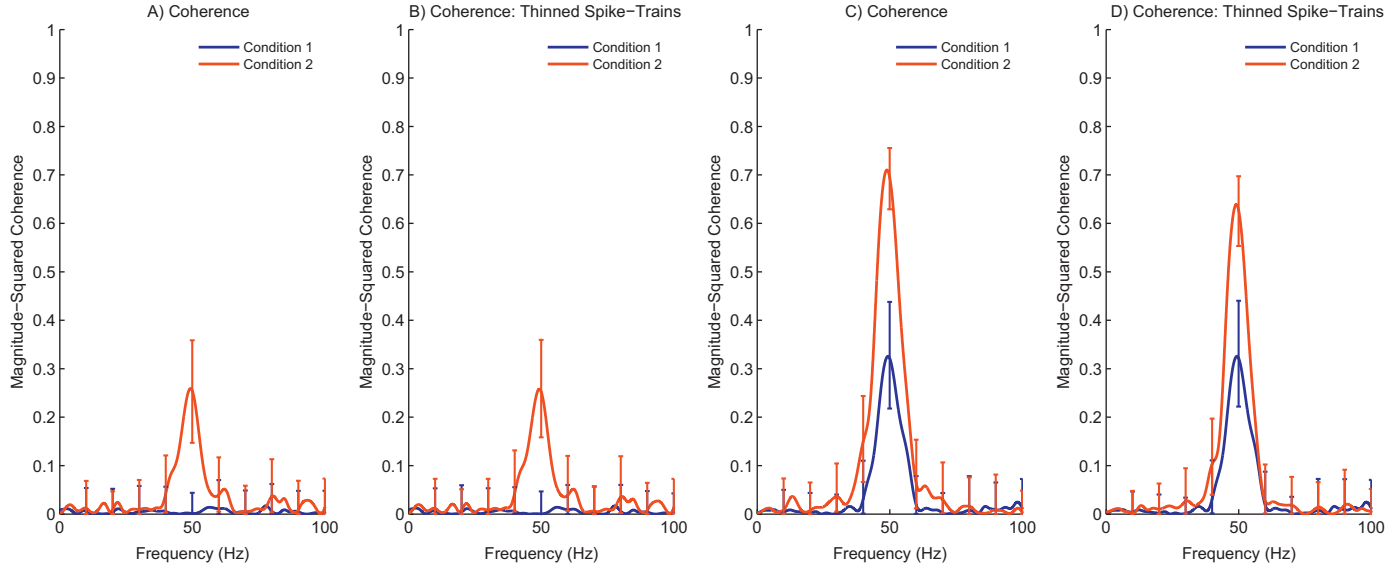

Fig. 6. Spike field coherence computed from the data generated in making (Fig. 4). *A*: Spike field coherence computed from the non-thinned spike-trains associated with the left-half of Fig. 4. *B*: Spike field coherence computed from the thinned spike-trains associated with the left-half of Fig. 4. *C*: Spike field coherence computed from the non-thinned spike-trains associated with the right-half of Fig. 4. *D*: Spike field coherence computed from the thinned spike-trains associated with the right-half of Fig. 4. The thinning procedure slightly reduces spike magnitude-squared coherence between the spikes and the LFP. Magnitude-squared coherence is more affected by the thinning procedure for the data generated with the log link function (C,D). This is due to the larger across-condition difference in the total number of spikes leading to larger probabilities of spike removal in the spike thinning procedure. The frequency resolution of these estimates, equal to twice the time-bandwidth parameter ($NW = 1$) divided by the observation duration of 200 ms, is 10 Hz.

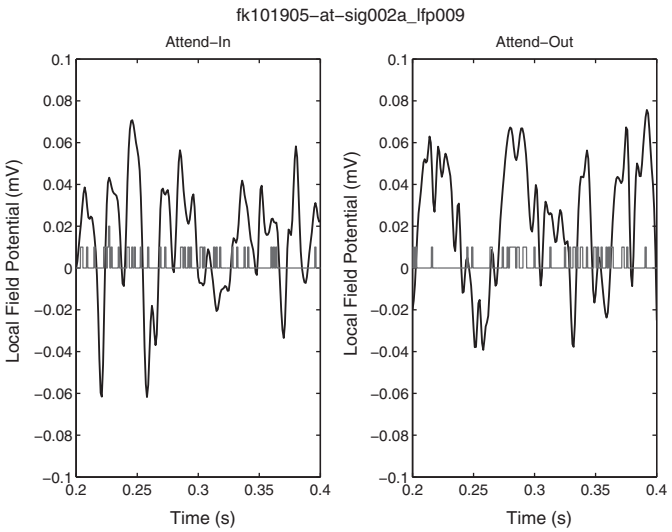

Fig. 7. Example local field potential (black curves) recorded from Macaque visual cortex during the attend-in (left) and attend-out (right) experimental conditions. Black vertical bars indicate the times at which spikes occur in multi-cell recordings from the FEF. The across-condition test for associativity between LFP rhythm and neural spike train is applied to time-series such as these.

Here y_t is the LFP and is taken to be a realization of an ARMA (5,2) process with spectrum depicted in Fig. 2, normalized such that $\max_t y_t = 1$. For the purposes of simulation, y_t is considered unitless. The coupling between the LFP and the rate, $\lambda_t^{(k)}$, is facilitated through the β parameter, which is equal to 80 Hz. The background rate, α_k, takes on the values: {60 Hz, 100 Hz, 140 Hz, 240 Hz}, depending on which simulation is performed (these rates map to the different quadrants on the left hand side of Fig. 3). An example of the rate computed using Eq. (29) is shown on the bottom left-hand side of Fig. 2. For each value of k, twenty trials each of one second duration are computed, each with $\lambda_t^{(k)}$ specified by Eq. (29). From this synthetic data, the squared-magnitude of the spike-field coherence, Eq. (6), is computed. The result is plotted in the left panels of Fig. 3. As expected, the squared-magnitude of the spike-field coherence decreases with increasing α_k. In the four plots on the right-hand-side of Fig. 3, the background rate, α, is held constant while the link between the LFP rhythm and the probability of spiking is varied. That is, instead of Eq. (29), the rate is specified as,

$$\lambda_t^{(k)} = \max(0, \tilde{\lambda}_t^{(k)}),$$
$$\tilde{\lambda}_t^{(k)} = \alpha + \beta_k y_t, \tag{30}$$

where $\alpha = 60$ Hz and β_k is equal to {80 Hz, 60 Hz, 40 Hz, 20 Hz} for $k = 1, 2, 3, 4$, when y_t is scaled such that $\max_t y_t = 1$. In both cases the squared-magnitude of the spike-field coherence decreases, demonstrating the confound between overall neural activity and the degree of association between LFP rhythm and spiking activity. On the right-hand side of six of the eight plots in Fig. 3, a bar indicating the p-values for the test introduced in Section 3 is presented. This test of the difference between LFP rhythm/spiking activity association, is applied pairwise between the $k = 1$ case and each of the cases indexed by $k = 2, 3$, and 4. The test correctly identifies changes in LFP rhythm modulated neural spiking probability; a feature that cannot be identified from either the spike-field coherence computed from the unthinned neural spiking events, or from the thinned neural spiking events (also plotted in Fig. 2).

In the second simulation, the difference between the piecewise linear link function and the log link function is highlighted. In this simulation, data are generated using the piecewise linear link function

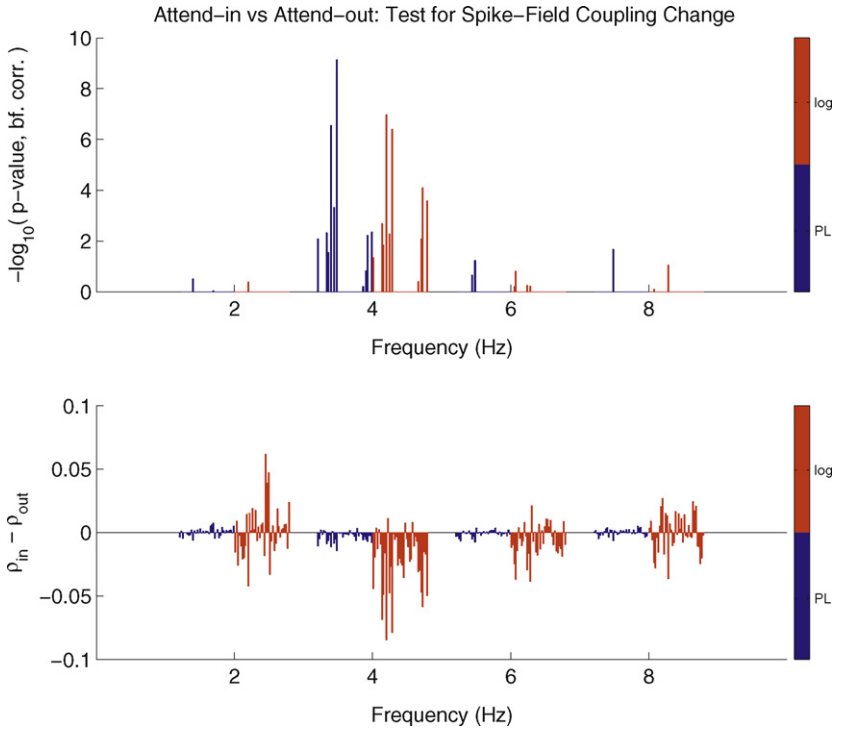

Fig. 8. Application of the proposed test (low-frequency) to real data. *Top*: The proposed test for spike-field coupling change applied, as a function of frequency, to LFP rhythm and neural spike trains recorded from one *Macaca mulatta* monkey trained in a covert attention task; as described in Gregoriou et al. (2009). Each frequency has 76 tests, two for each spike-train/LFP pair. Each test was computed using both the PL link function (blue) and the log link function (red). Each test is sensitive to a change across attentional condition but in different ways. The tests have been Bonferroni corrected for multiple comparisons. *Bottom:* Difference in the modulation, ρ, across attention conditions. At 4 Hz, spike coupling to LFP rhythm is reduced during attention and is accompanied by a spike-phase distribution that is less tuned to a preferred phase of spiking. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

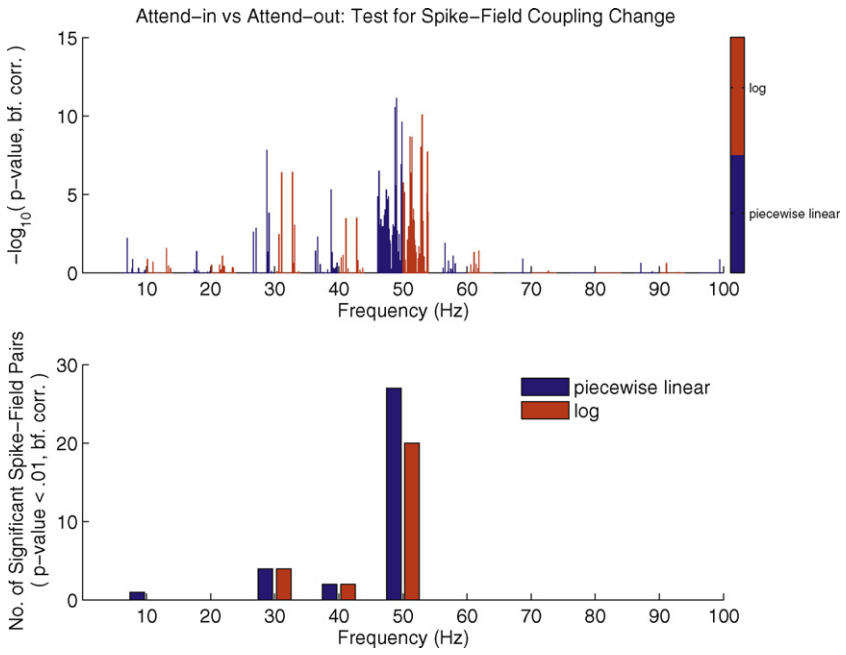

Fig. 9. Application of the proposed test (high frequency) to real data. *Top*: The proposed test for spike-field coupling change applied, as a function of frequency, to LFP rhythm and neural spike trains recorded from one *Macaca mulatta* monkey trained in a covert attention task; as described in Gregoriou et al. (2009). Each frequency has 76 tests, two for each spike-train/LFP pair. Each test was computed using both the PL link function (blue) and the log link function (red). Each test is sensitive to a change across attentional condition but in different ways. The tests have been Bonferroni corrected for multiple comparisons. *Bottom*: Number of pairs with significant differences across conditions vs. frequency. Most changes in association occur for a frequency interval centered upon 50 Hz. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

(left-hand side of Fig. 4), and using the log link function (right-hand side of Fig. 4). Here the LFP is as in the first simulation and the modulation, ρ, for the data generated by the PL link, is changed across conditions while the background α is held constant. Specifically, $\alpha = 40$ Hz, and the modulation varies from 0 Hz to 40 Hz. Note that the apparent modulation of 20 Hz in Fig. 4(A) is due to the variability of the autoregressive LFP. When the data are generated using the piecewise linear link the probability densities of the LFP phase at spiking times differ between conditions, Fig. 4(B), and the tests associated with each of the two models show significant differences between the two conditions, Fig. 4(D). For the data generated using the log link, the modulation, ρ, is held constant while the background, α, is changed. Specifically, ρ is set to 1.3 and *alpha* varies from 3.0 for condition 1 to 4.4 for condition 2. In this case, the probability of the phase at which spikes occur has a constant tuning width (Fig. 4(F)), but the rate is modulated in the PL sense more strongly by the LFP rhythm. Compare with Fig. 4(E). Only the test computed using the piecewise linear link function show significant differences between the two conditions, Fig. 4(F). The test computed using the log-link is designed not to respond to changes in coupling that do not affect the distribution of the phase at which spikes occur. In particular, the relationship of the LFP rhythm with the spiking rate is not linear and involves sinusoids at frequencies outside of the putative LFP rhythm in the process of spike train generation. This nonlinearity allows for a "tuning" of spike times to a specific phase of the LFP. More specifically, the modulation in the log-link model controls the width of the spike-time LFP phase density about the preferred phase of spiking. Because of this, the test procedure using the log link model allows for comparisons in the width of phase tuning. Note that standard procedures involving spike-field coherence are not sensitive to phase tuning changes.[9]

In Fig. 5, the modulation, *rho*, estimated from the data depicted in Fig. 4 is shown. The modulation is consistent with the results of the proposed hypothesis test plotted in Fig. 4(D) and in Fig. 4(H). Associated with Figs. 4 and 5 are plots of magnitude-squared spike-field coherence. These are presented in Fig. 6. In these two cases (left and right halves of Fig. 4), spike-field coherence responds to the across-condition changes in a fashion similar to the across-condition changes detected by the proposed test associated with the PL link function, but does not respond in a fashion similar to the response of the proposed test associated with the log link function. This similarity to the behaviour of the test associated with the PL link function is superficial, in that spike-field coherence and the proposed test respond differently in the first simulation, see Fig. 3.

5. Data

The across-condition test for associativity between LFP rhythm and neural spike train is applied to signals recorded from one *Macaca mulatta* monkey trained in a covert attention task as previously described in Gregoriou et al. (2009). Briefly, the monkeys task was to detect a color change of a target stimulus presented among distracters. The targets location was randomized in different trials so that attention could be directed inside ("attend-in") or outside ("attend-out") the receptive-field (RF) of the recorded neurons. Spike trains corresponding to multi-unit activity as well as LFPs were recorded simultaneously from the frontal eye fields (FEF) and visual area V4. Spike trains were obtained after filtering the recorded signal between 250 Hz and 8 kHz, and amplifying and digitizing the signal at 40 kHz. The LFP signals were obtained after filtering between 0.7 Hz and 170 Hz, and amplifying and digitizing at 1 kHz. LFP data were post-processed to correct for the known phase shifts as previously described in Gregoriou et al. (2009). Fig. 7 shows an example LFP time series and associated spike train raster for the attend-in and attend-out experimental conditions. Data from 3 sessions including 12 spike signals and

[9] That is, the width of the phase distribution can change without altering the spike-field coherence.

11 LFP signals were used in this study. The resulting 38 spike-LFP pairs recorded on different electrodes were submitted to the test for spike-LFP coupling. In a first, low frequency analysis, the test is applied from 2 Hz to 8 Hz with a bandwidth of 2 Hz, and reveals significantly different associativity at a frequency of 4 Hz, as shown in Fig. 8, top. This change in associativity results in reduced LFP-neural modulation when attending (Fig. 8, bottom), and is accompanied by a reduced locking of the 4 Hz LFP phase to a preferred phase of spiking. This former inference is made by examining the test results and associated modulation change when using the PL link, and the latter inference is made using the log link.

The test is applied from 5 Hz to Nyquist frequency – 5 Hz with a bandwidth of 10 Hz, and reveals patterns of significantly different associativity between different LFPs and spike trains at many frequencies, as shown in Fig. 9. In particular, one sees that the tests associated with either link function are significant for many spike-train/LFP pairs at 50 Hz, with some highly significant across-condition changes in spike-train/LFP rhythm coupling occurring at frequencies below 50 Hz. In the remainder of this work a frequency of 50 Hz is focused upon due to the large number of significant tests at this frequency. This choice is consistent with the frequency interval where significant differences in coherence were found with attention in Gregoriou et al. (2009). The per-frequency modulation associated with the PL link model and the log link model is plotted in Fig. 10 alongside magnitude-squared coherence. Modulation is larger in the attend-in condition and at frequencies below 60 Hz. It increases monotonically with frequency from 0 Hz to 50 Hz. These changes in modulation occur whether using the PL link or the log link. This indicates, while the monkey attends, an increased spiking at frequencies less than or equal to 50 Hz, accompanied by an increase in the degree of phase locking to a specific preferred-LFP-phase of firing. Referring to the discussions in Sections 2.1 and 3, and to the simulation examples in Section 4, increased 50 Hz firing is consistent with the interpretation of modulation associated with the PL link while tighter phase locking is consistent with the interpretation of modulation associated with the log link. Measures

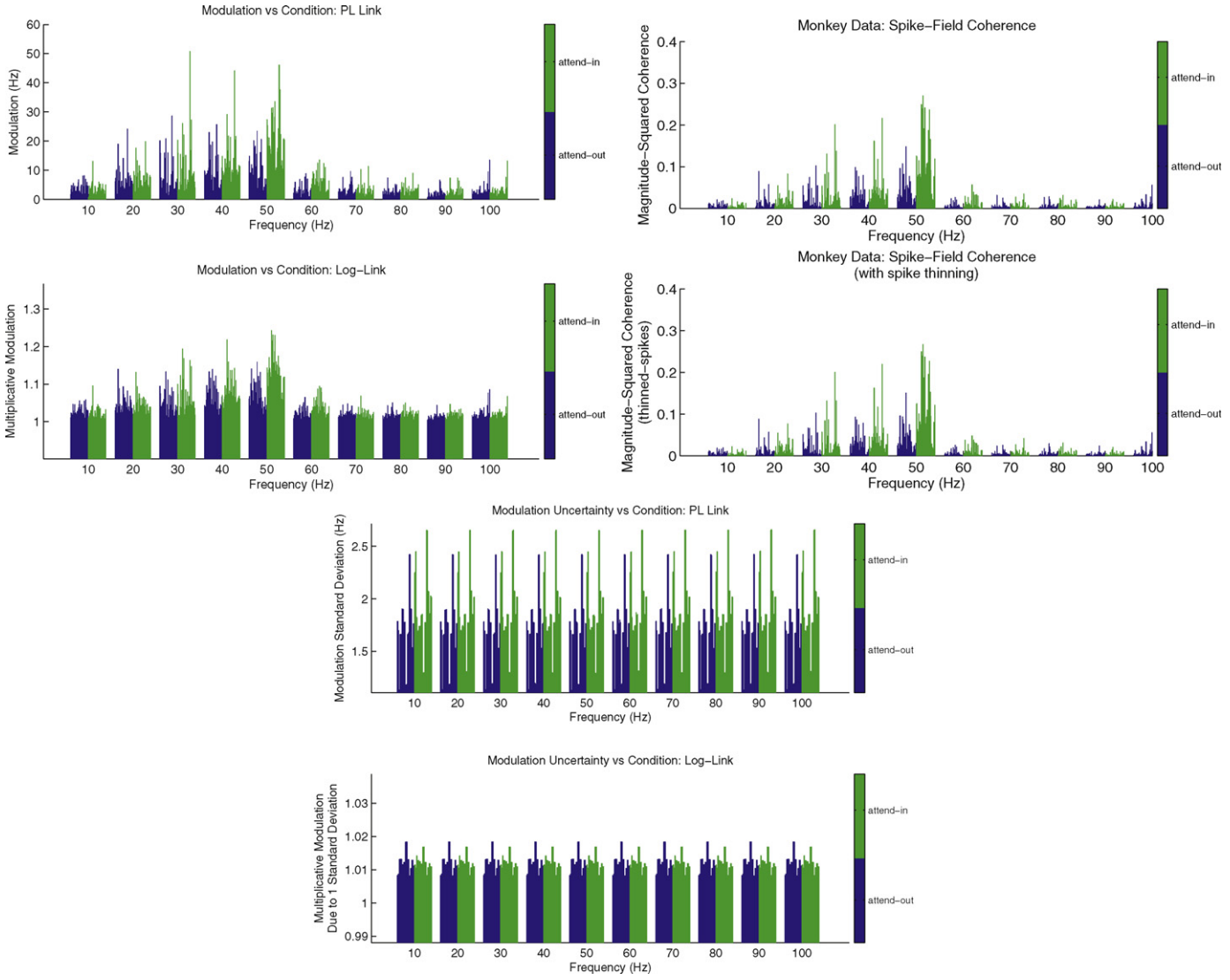

Fig. 10. Modulation vs. experimental context for the Macaca monkey data as a function of frequency. The modulation is computed using the PL model (left, top), and using the log link model (left-bottom). Associated modulation uncertainty is plotted below, center. Magnitude-squared coherence (MSC) (right-top) and MSC computed from thinned spike-trains (right-bottom), behave in a similar fashion to modulation. In all cases, modulation (PL link or log link) and MSC behave similarly; during the attend-in condition it is larger than during the attend-out condition, and this is true for almost all frequencies. In addition, there is a linearly increasing trend from 0 Hz to 50 Hz followed by a sharp decrease. In all plots, each bar, for a given condition, represents the modulation, MSC, or modulation uncertainty for a single LFP/electrode pair for a specific experimental condition. The ordering of the bars is consistent from one condition to the other so that direct comparisons across condition can be made visually. Due to the large, spiking rates, and the small difference in spiking rates across condition, the MSC computed from the thinned and the non-thinned spike trains is nearly indistinguishable.

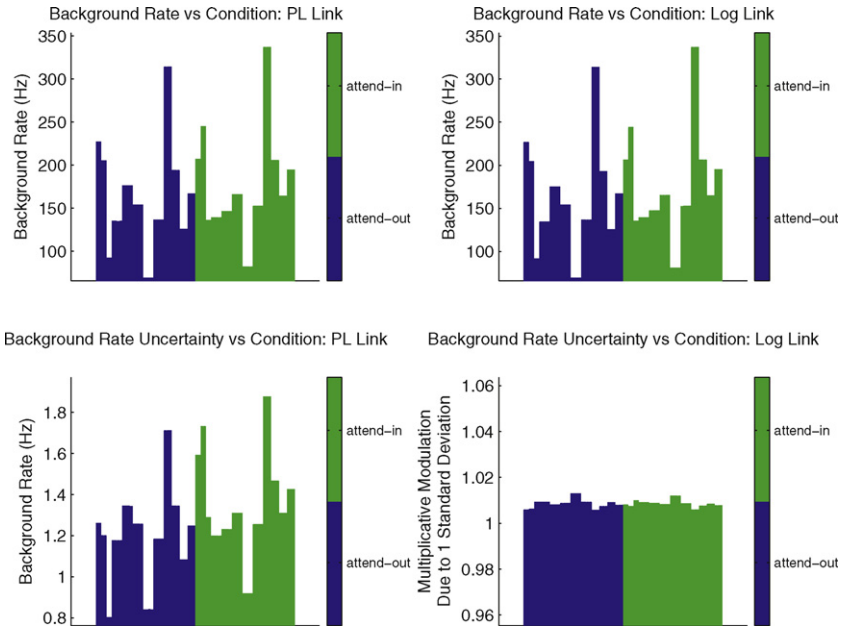

Fig. 11. Background rates (α): The height of each bar of the same color indicates the estimated background rate for one LFP/spike-train pair. There are 38 such pairs and the color indicates the experimental condition. The background rate estimates, computed with either the PL link function or the log link function are, as plotted, indistinguishable. Uncertainty is larger for larger rates, a property expected of count-type data. Background rates tend to be larger during the attend-in experimental condition. The uncertainty associated with the log link is e^{σ_α}; specifying the multiplicative-modulation of the background rate resulting from one positive standard deviation in *alpha* when using the log link function.

of uncertainty associated with the modulation are also plotted in Fig. 10. On this scale, the uncertainty is independent of frequency. Its computation is provided in Appendix D. In Fig. 11, the estimated α parameters for the data are plotted in terms of rate modulation. When using the PL link the estimated background rate is α and when using the log link the estimated background rate is e^{α}. Both of these quantities are multiplied by Δ^{-1} to convert from spikes per bin to spikes per second (i.e. Hz). The background rates estimated using the different link functions are identical on the scale plotted in Fig. 11.

In Fig. 12, the across-condition average spiking rate difference is plotted against the average across condition spiking rate. The average rates are large and the average rate difference across conditions is typically an order of magnitude smaller, see Fig. 12. Because of this thinning procedures are expected to have a modest effect upon magnitude-squared coherence.

In Fig. 13, the proposed tests are seen to be strongly correlated. The test associated with the piecewise-linear link function is strongly correlated with the magnitude-square coherence difference; with a different offset associated with differing across-condition average rate differences demonstrating the response of magnitude square coherence difference to average rate difference. The strong linear relation between the magnitude-square coherence and the significance of the proposed test employing the piecewise-linear link function demonstrates the similarity between these two measures: both tests are responding to across-condition changes (other than changes in the average rate) in a similar way. The $-\log_{10}(p-\text{values})$ associated with the test using the log link and the piecewise-linear link are correlated with smaller p-values tending to be associated with the PL link function (Fig. 13, right); but with both tests reporting detections and misses absent from the other. That is, for most spike train/LFP pairs, the test results suggest that changes in attention are associated with both a change in the spike train/LFP coupling strength at 50 Hz, but also with a change in the distribution of the LFP phase at the time of spiking. Caution is required in this interpretation because, as previously mentioned in Section 3, the model using the log link function is potentially biased by activity at frequencies adjacent to 50 Hz. The behaviour of the tests is further illustrated in Fig. 14, where a raster of significant detections of across-condition change is depicted. In this raster, each dot corresponds to a significant test. Significant detections occur much more frequently with the proposed testing procedures than with a test based upon the bootstrap confidence intervals of magnitude squared coherence. While there is much overlap between the two different types of tests, some

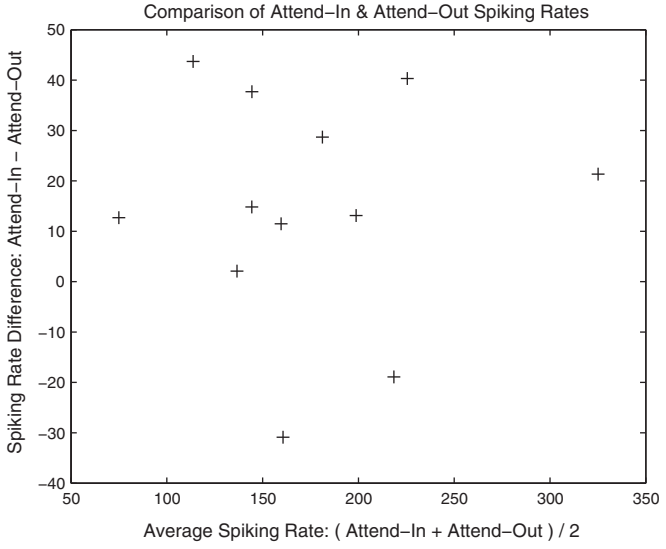

Fig. 12. Difference of the across-condition average spiking rates plotted against the average of the across-condition average spiking rates. The average spiking rates tend to be large while the across-condition rate changes are approximately an order of magnitude less. The thinning operation has a no effect due to the typical pair-wise (i.e. across-condition) similarity of the spiking rates (Fig. 12); as well as the relatively large firing rates.

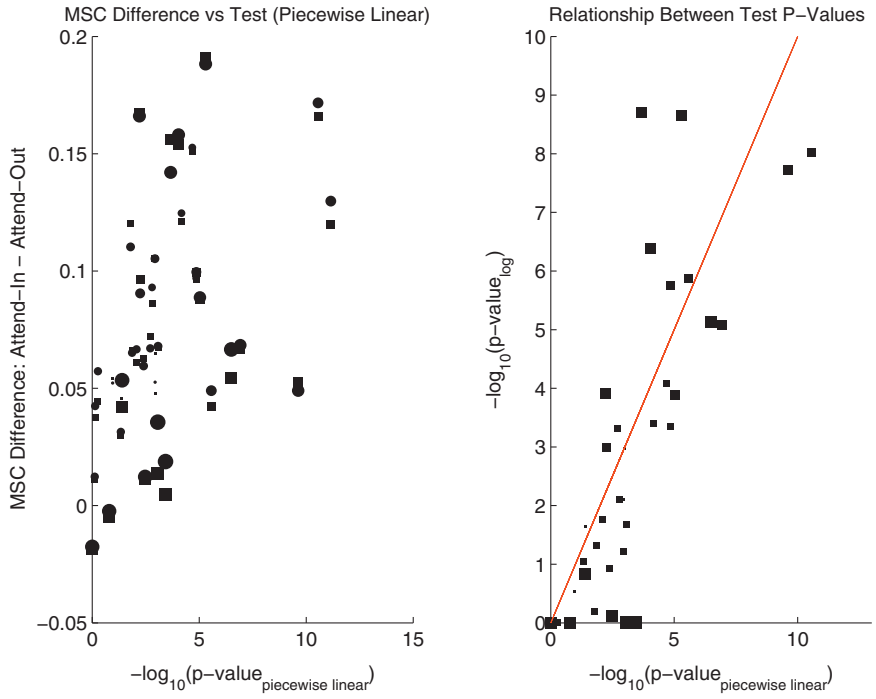

Fig. 13. Test comparisons. Summary: all three tests are correlated. The thinning procedure tends to reduce MSC difference across-condition. There exist a number of significant tests associated with the PL link function, that do not have significant counter-parts when using the log link function. This latter fact indicates that for some LFP/spike-train pairs a distributional change in the LFP phase of spiking is not accompanying a change in spike-field association. *Left*: The difference between the magnitude-squared coherence evaluated at 50 Hz for the attend-in condition is reduced by the magnitude-squared coherence at 50 Hz for the attend-out condition and is plotted against the $-\log_{10}(p-\text{values})$ computed using the proposed testing procedure with the model employing the piecewise linear link function. Circles indicate the use of non-thinned spike-trains and squares indicate that computations have occurred using spike-trains thinned to the minimum average firing rate of the two spike-trains. Each shape corresponds to a single spike-train/LFP pair and the size of each shape is proportional to the difference in the across-condition average spiking rates. *Right*: Comparison of *p*-values computed using the proposed testing procedure with differing functions linking the expected intensity to the covariates. A linear trend relating the transformed test *p*-values exists; with the test computed using the log link tending to have smaller *p*-values.

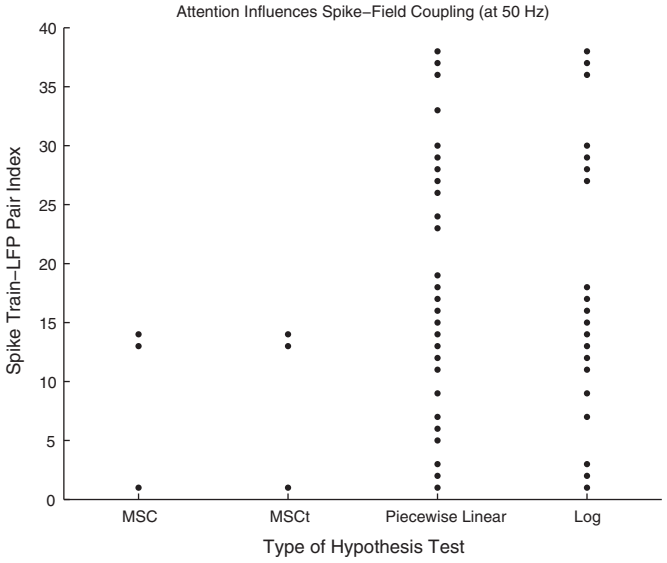

Fig. 14. A raster of significant detections. *MSC* refers to tests performed by comparing 99% magnitude-squared coherence bootstrap confidence intervals across conditions (a detection corresponds to non-overlapping confidence intervals), *MSCt* refers to the same test but computed using spike-trains randomly thinned to the minimum pairwise average spiking rate; while *Log* and *Piecewise Linear* refer to the proposed test computed using the different link functions. The *MSC* and *MSCt* detections are detected by both of the proposed tests and the proposed tests are similar with the exception of a few spike-train/LFP pairs. Here a significant *p*-value is taken to be 0.01 after Bonferroni correction. There are no tests computed using the log link function that are significant where the associated test computed using the PL link function is not significant. Tests computed using the PL link function that are significant where the associated test computed using the log link function is not significant indicate changes in the modulation due to LFP at 50 Hz that do not change the width of the distribution of the LFP phase associated with spikes.

detections present using the test computed with the piecewise linear link function are not present when computing the test using the log link function.

The proposed tests detect all of the across-condition changes detected by the magnitude-squared coherence test as well as many that are not detected by the magnitude-squared coherence test. One expects that the proposed tests based upon parametric modelling will possess greater statistical power than the test based upon the magnitude-squared coherence bootstrap confidence intervals. There are two reasons for this. First, the proposed tests are comparing to a *p*-value (prior to corrections for multiple comparisons) of 0.05 under exact asymptotic distributions whereas the test based upon the magnitude-squared coherence bootstrap confidence intervals is significant when the confidence intervals do not overlap. Second, the estimators $\hat{\beta}_c$ and $\hat{\beta}_s$ are attaining, asymptotically, the Cramér-Rao lower bound on estimator variance (see Fig. 15 in Appendix B, Scharf (1991, p. 221) and Hogg et al. (2005, p. 325)), while the bootstrap confidence intervals do not satisfy such an optimality property. A more thorough study of testing performance is left for future work.

6. Discussion

A testing procedure capable of disambiguating the effect of frequency dependent association from the effect of changes of average neural spiking rate on a commonly employed measure of spike field coherence is presented. This method, described in Section 3, exploits a generalized linear model of neural spiking activity to arrive at a statistical hypothesis test independent of the effect of background rate. The method, employing differing link functions is demonstrated on synthetic data in Section 4. The proposed testing procedure is shown to respond to different features of the data

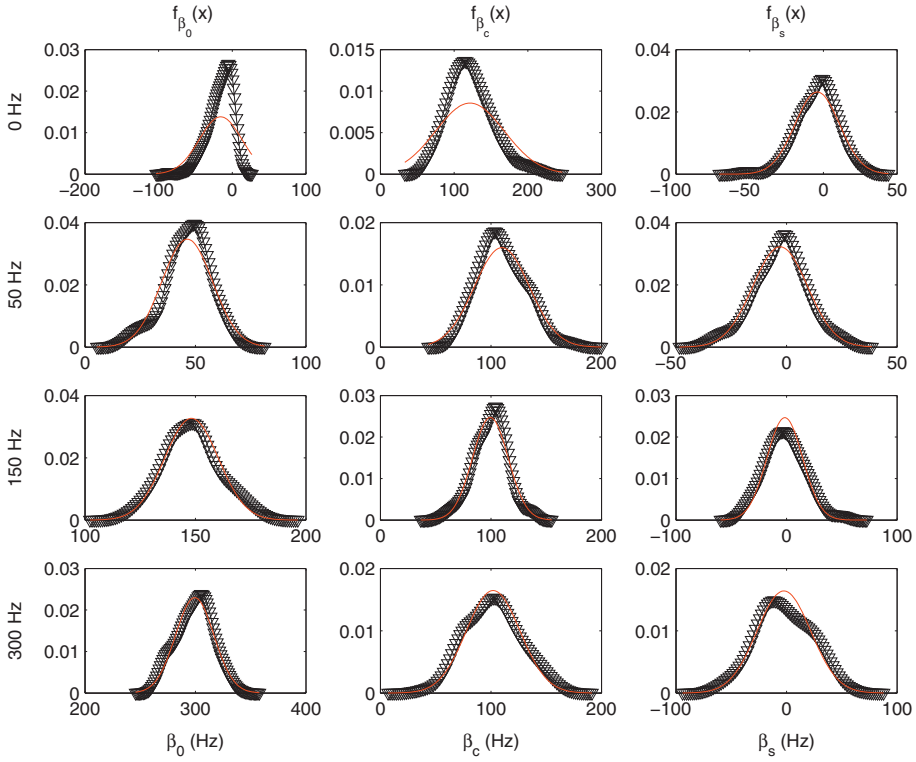

Fig. 15. Black curve: empirical probability density functions for $\hat{\beta}_0$ (left column), $\hat{\beta}_c$ (middle column) and $\hat{\beta}_s$ (right column), for background rates increasing from top row to bottom row. Empirical probability density functions are computed from 150 estimates. Each estimate is computed from a single trial consisting of a one second recording generated according to Eqs. (8) and (9), with a background rate specified by the left-hand-side label for each row of the figure, and with a constant LFP-rate coupling constant equal to 100 Hz. The Gaussian approximation is good for all rates and is excellent for rates greater than or equal to 50 Hz. The non-linear link function is in a linear regime for background rates exceeding 100 Hz. As expected, in this regime estimator variance increases with background rate. When the background rate is 0 Hz, the rate is a rectified cosine, and though bias is evident; the estimates are sufficiently accurate to perform inference.

depending on the link function employed both in simulation and on real data; allowing for the investigation of both a per-frequency change in spike-train/LFP association and the effect of average spiking rate upon spike-triggered phase distributions. Table 2 provides a summary of the performance of the three tests for a specific type of across-condition change.

Due to the parametric nature of the proposed methodology, the testing procedure allows for the inclusion of any covariates of interest, and the full statistical machinery associated with the point process generalized linear spike-train modeling framework is available (Truccolo et al., 2005). While the generalized linear model employing the piecewise linear function of the covariates of interest is resilient to unmodelled sinusoids when average spiking rates are sufficiently large relative to the size of the modulation, ρ, when average spiking rates are lower or when the log link function is employed, these unmodelled sinusoids may be important. In this situation, available model selection procedures will be important.

In the proposed methodology history dependence is explicitly modelled. One notes that if the form of the history dependence is oscillatory, identifiability problems may arise. Because the conditional log-likelihood is convex for all parameter choices, this non-identifiability will manifest in a flat likelihood and parameter estimates will be associated with large variances. Thus, the procedure is robust to non-identifiability in the sense that this problem will be transparent to the analyst.

The procedures detailed in this paper require the estimation of the instantaneous phase of the local field potential for relatively small frequency intervals. In this estimation, care must be taken to ensure that the phase estimated for any given interval of frequencies is not due to large signal in an adjacent frequency interval that is "leaking" into the interval for which phase is being estimated. Further, care must be taken to appropriately discard data where bandpass filtering has introduced an edge-effect. This edge-effect will depend both upon the type and order of the filter employed.

In Vinck et al. (2010, 2011) the pairwise phase consistency (PPC) approach to studying spike-field association is introduced. These estimators, introduced to remove amplitude dependence and reduce bias, provide an interesting alternative to the proposed

Table 2
Summary of test properties. When considering changes, it is assumed that all parameters besides the changing parameter are fixed in the data generating model. "Change in PL α" refers to a change in the α parameter when the data is generated by the model, Eqs. (8) and (9), when using the piecewise linear link function. Similarly, "Change in LOG ρ" refers to a change in the ρ parameter when the data is generated according to the model using the log link function.

Property	PL link	Log link	Spike-field coherence
Affected by unmodelled sinusoids	NO[a]	YES	NO
Affected by change in PL α	NO	YES	YES
Affected by change in PL ρ	YES	YES	YES
Affected by change in Log α	YES	NO	YES
Affected by change in Log ρ	YES	YES	YES

[a] When the background rate sufficiently large relative to the combined oscillatory activity. See Section 2.2.

procedure. In the current work a testing procedure based on asymptotic distributions of relevant estimators is used to provide tests for across-condition change in spike-field association, while equivalent tests involving PPC have not yet been documented. A complete comparison of the two statistical procedures is left for future study.

6.1. Rate-free spike-field coherence estimation

In Eq. (7), it is seen that the spike-field coherence, $C_{ny}(f)$, is related to the intensity-field coherence, $C_{\Lambda y}(f)$. This latter quantity directly relates the probability of spiking to the local field potential rhythm, independent of the average neural rate. That is, the theoretical quantity to be estimated, in this case the intensity-field coherence, $C_{\Lambda y}$, does not depend on the average neural rate. Critically this differs from having an estimate of the intensity-field coherence with an accuracy that does not depend on average neural rate. In particular, if the majority of spikes are random and not due to LFP rhythm influence, one expects the ability to determine the intensity-field coherence to deteriorate. In effect, the signal to noise ratio is reduced. The estimated intensity-field coherence, calculated from observed data, will have an accuracy dependent upon the degree to which a rhythm in the local field potential affects the spike times relative to the effect upon spike times by contributions unassociated with the local field potential rhythm. These contributions include, for example, contributions to the spiking probability from the influence of past spiking activity, contributions from the influence of uninteresting covariates, or an elevated probability of spiking due to an elevated background firing rate. In essence, the larger the rate for a given level of association, the greater the "noise" obscuring the association. Thus, while "rate-free" quantities like the intensity-field coherence, $C_{\Lambda y}(f)$, exist, the variability (and hence the typical accuracy) of such estimators necessarily depends upon overall neural activity. While a fact of life, this does not mean that quantities such as the intensity-field coherence are useless; but, rather, that a principled statistical procedure involving the intensity-field coherence will account for sampling properties that will detail rate-dependent accuracy. In this paper, by employing the existing point process methodology introduced in Truccolo et al. (2005), a statistically principled parametric modelling approach is taken to perform between-condition comparison of spike-train LFP coupling. This procedure explicitly accounts for changes in firing rate across-condition and separates effects of these changes from changes in the modulation of spiking activity associated with local field potential rhythm. Unlike current procedures associated with spike-field coherence, the method accurately and explicitly assesses the uncertainty of relevant parameter estimates and appropriately deals with changes in signal-to-noise ratio.

Acknowledgements

KQL acknowledges NSF award, [DMS-1042134]. UTE acknowledges support from [IIS 0643995] and [DMS 1042134]. UTE and MAK acknowledge support from the National Institute of Neurological Disorders And Stroke, Award Number [R01NS072023]. SG was supported by the National Institute of Mental Health, NIH, Division of Intramural Research. This work was partially supported by the National Institute of Neurological Disorders and Stroke [R01NS073118] and [R01NS072023], as well as by grants, [EY017292] and [EY017921].

Appendix A. Maximum likelihood estimate

Let dn_j be the jth increment of a discrete-time point-process modeling neuron spiking behavior, and define the associated stochastic conditional intensity, Λ_j, as described in Section 2, according to Eqs. (8) and (9). Further, model the field-type time series as a truncated realization of the random process, y_j, at time-index j. For convenience, define the vectors, $\boldsymbol{\beta} = [\beta_0\, \beta_c\, \beta_s]^T$, $\mathbf{dn} = [dn_1\, dn_2 \ldots dn_n]^T$, and $\mathbf{y} = [y_1\, y_2 \ldots y_n]^T$, where ^{T} denotes matrix transposition. The full likelihood is

$$L(\boldsymbol{\beta}, \boldsymbol{\phi}) = f(\mathbf{dn}, \mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\phi}), \tag{31}$$

$$= f(\mathbf{dn}|\boldsymbol{\beta}, \boldsymbol{\phi}, \mathbf{y}) f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\phi}), \tag{32}$$

$$= f(\mathbf{dn}|\boldsymbol{\beta}, \boldsymbol{\phi}) f(\mathbf{y}|\boldsymbol{\phi}). \tag{33}$$

Thus, the full likelihood for both the LFP to conditional intensity coupling parameters, $\boldsymbol{\beta}$, and the instantaneous phase of the local field potential, $\boldsymbol{\phi}$, factors into the product of two terms. The first term is the likelihood in the GLM cosine-tuning model (when considering $\boldsymbol{\phi}$ to be known), and the second term is the contribution to the full likelihood due to the probability of the observed local field potential given knowledge of the instantaneous phase of the local field potential. The log-likelihood, $\ell(\boldsymbol{\beta}, \boldsymbol{\phi})$ is,

$$\ell(\boldsymbol{\lambda}(\boldsymbol{\beta}, \boldsymbol{\phi}), \boldsymbol{\phi}) = \sum_{j=1}^{N} dn_j\, \log(\Delta\lambda_j) - \Delta\lambda_j - \log(dn_j!) + \log(f(\mathbf{y}|\boldsymbol{\phi})) \tag{34}$$

where, on the right-hand-side of Eq. (34), the explicit dependence of the conditional intensity, λ_j at time-index j, on the vector of parameters, $\boldsymbol{\beta}$, is suppressed. The first two terms in Eq. (34) are obtained from $f(\mathbf{dn}|\boldsymbol{\beta}, \phi)$ by conditioning and then marginalizing out the stochastic conditional intensity within a recursive operation. In the following the notation, dn_t, is reserved for the random quantity who's realizations are used to model the observed counts n_t at time index t. This more explicit notation, in contrast to that used in Eq. (33), is adopted for clarity. That is,

$$\begin{aligned} f_{dn_N, dn_{N-1}, \ldots, dn_1}(n_N, n_{N-1}, \ldots, n_1) &= \int_{-\infty}^{\infty} f_{dn_N|H_N, \Lambda_N}(n_N|\check{\lambda}_N, h_n) f_{\Lambda_N, H_N}(\check{\lambda}_N, h_N) d\check{\lambda}_N, \\ &= \int_{-\infty}^{\infty} \frac{(\Delta\check{\lambda}_N)^{n_N} e^{-\Delta\check{\lambda}_N}}{n_N!} f_{\Lambda_N, H_N}(\check{\lambda}_N, h_N) d\check{\lambda}_N, \end{aligned} \tag{35}$$

where H_N is the collection of random variables representing spike counts occurring prior to time index N, i.e. $H_N = \{dn_{N-1}, dn_{N-2}, \ldots dn_1\}$. Associated with these random variables are their realizations, h_N, the collection of counts occurring prior to time-index N. Continuing, consider the joint density of the stochastic conditional intensity with the process history, H_N,

$$f_{\Lambda_N, H_N}(\check{\lambda}_N, h_N) = f_{\Lambda_N|H_N}(\check{\lambda}_N|h_N) f_{H_N}(h_N), = \delta(\check{\lambda}_N - \lambda_N) f_{H_N}(h_N), \tag{36}$$

where δ is the Dirac delta function, and λ_N is specified according to Eqs. (8) and (9). In particular,

$$\lambda_N = \max(\tilde{\lambda}_N, 0), \tag{37}$$

and

$$\tilde{\lambda}_N = \alpha + \beta_c \cos(\hat{\phi}_N) + \beta_s \sin(\hat{\phi}_N) + \sum_{k=1}^{K} \gamma_k dn_{N-k}. \tag{38}$$

Combining Eq. (38), Eq. (36) and Eq. (35), one obtains the recursive relation for the joint density of all of the spike counts in terms of a term given by the likelihood multiplied by the joint probability

mass function of all of the spike counts less the count for the last bin:

$$\begin{aligned} f_{dn_N,dn_{N-1},\ldots,dn_1}(n_N,n_{N-1},\ldots,n_1) &= \frac{(\Delta\lambda_N)^{n_N}e^{-\Delta\lambda_N}}{n_N!}f_{H_N}(h_N),\\ &= \frac{(\Delta\lambda_N)^{n_N}e^{-\Delta\lambda_N}}{n_N!}\times f_{dn_{N-1},\ldots,dn_1}(n_{N-1},n_{N-2},\ldots,n_1),\\ &= \prod_{k=1}^{N}\frac{(\Delta\lambda_k)^{n_k}e^{-\Delta\lambda_k}}{n_k!}. \end{aligned} \tag{39}$$

Here the likelihood of the spike counts conditioned on the LFP instantaneous phase estimates is written in Eq. (39) by completing the recursion. Reverting back to our previous notation, Eq. (39) becomes,

$$\prod_{k=1}^{N}\frac{(\Delta\lambda_k)^{dn_k}e^{-\Delta\lambda_k}}{dn_k!}, \tag{40}$$

where the spike count at time index k, referred to as dn_k in Eq. (34), and n_k in Eq. (39) is once again referred to as dn_k. This latter notation is used throughout the following.

Due to monotonicity, the location of the maximum of the likelihood and of the log-likelihood are identical. Computing first derivatives one obtains the score equations (Casella and Berger, 2001),

$$\frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\phi})}{\partial\beta_j} = \sum_{t=1}^{N}\frac{dn_t}{\lambda_t}\frac{\partial\lambda_t}{\partial\beta_j} - \Delta\sum_{t=1}^{N}\frac{\partial\lambda_t}{\partial\beta_j}, \tag{41}$$

since the second and fourth terms are independent of $\boldsymbol{\beta}$. Similarly,

$$\frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\phi})}{\partial\phi_j} = \frac{1}{f(\mathbf{y}|\boldsymbol{\phi})}\frac{\partial f(\mathbf{y}|\boldsymbol{\phi})}{\partial\phi_j}. \tag{42}$$

The derivative with respect to the conditional intensity, λ_t, is

$$\frac{\partial\lambda_t}{\partial\beta_j} = \begin{cases} \dfrac{\partial\tilde{\lambda}_t}{\partial\beta_j}, & \tilde{\lambda}_t > 0\\ 0, & \tilde{\lambda}_t < 0\\ \text{undefined}, & \tilde{\lambda}_t = 0 \end{cases}, \tag{43}$$

and, when $\tilde{\lambda}_t > 0$,

$$\frac{\partial\lambda_t}{\partial\beta_j} = \begin{cases} 1, & j = 0\\ \cos(\phi_t), & j = c\\ \sin(\phi_t), & j = s \end{cases}, \tag{44}$$

where the indices, $j = \{0, c, s\}$, are used, consistent with the indices employed in Section 2. One notes that there will be instances when $\tilde{\lambda}_k = 0$ such that $\partial\lambda_t/\partial\beta_j$ will be defined. An example is when $\beta_j = \beta_s$ and $\phi_t = 0$. These isolated instances are the exception rather than the rule and are ignored in Eq. (43) and in the following. The negative of the Hessian, or second derivative with respect to the parameters, is equal to the observed Fisher information matrix, $\mathbf{I}_\beta$, discussed in Section 2. As in Eq. (43), the second partial derivative of the log-likelihood with respect to β_k is non-trivial and defined when $\tilde{\lambda}_t > 0$. Then,

$$\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\phi})}{\partial\beta_j\partial\beta_k} = -\sum_{t=1}^{n}\frac{dn_t}{\lambda_t^2}\frac{\partial\lambda_t}{\partial\beta_j}\frac{\partial\lambda_t}{\partial\beta_k}, \quad \tilde{\lambda}_t > 0, \tag{45}$$

since $\partial^2\tilde{\lambda}_t/\partial\beta_j\partial\beta_k = 0$ for all of the indices, t, j, and k. More succinctly, Eq. (45), can be written

$$\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\phi})}{\partial\beta_j\partial\beta_k} = -(\mathbf{H}^T\mathbf{D}\mathbf{H})_{j,k}, \quad \tilde{\lambda}_t > 0, \tag{46}$$

where, as defined in Section 2, $\mathbf{H} = [\mathbf{1}\,\mathbf{c}\,\mathbf{s}]$. Here $\mathbf{1}$ is an n-element column vector of ones and the tth element of the column vectors $\mathbf{c}$ and $\mathbf{s}$ is, respectively, $\cos(\phi_t)$ and $\sin(\phi_t)$. The diagonal matrix, $\mathbf{D}$, is defined as,

$$(\mathbf{D})_{j,j} = \frac{dn_j}{\lambda_j^2}. \tag{47}$$

Note that $\mathbf{H}^T\mathbf{H}$ is approximately diagonal due to the orthogonality of sinusoids at different frequencies. [10] Then $\frac{dn_j}{\lambda_j^2}$ is proportional to, and of the same sign as, the eigenvalues of $\mathbf{H}^T\mathbf{D}\mathbf{H}$. Hence the Hessian is negative definite and the log-likelihood is convex in the unknown parameters $\boldsymbol{\beta}$; implying a unique maximum for the likelihood as a function of $\boldsymbol{\beta}$.

The $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ which root Eq. (41) are the maximum-likelihood estimators, yielding the background rate, α, the strength of coupling, ρ, and the preferred phase of coupling, ϕ_p, between the LFP rhythm and the spike times, as well as the instantaneous phase, $\boldsymbol{\phi}$, of the LFP at the frequency of interest, f_0. As described in Section 2, in this work, the maximum-likelihood estimator of $\boldsymbol{\phi}$ is approximated by the Hilbert transform type estimator, $\hat{\boldsymbol{\phi}}$. This affects both the maximum likelihood estimator of $\boldsymbol{\beta}$ through the dependence of $\boldsymbol{\lambda}(\boldsymbol{\beta},\boldsymbol{\phi})$ on the LFP rhythm instantaneous phase, $\boldsymbol{\phi}$, as well as the curvature of the log-likelihood evaluated at the maximum likelihood estimates. Hence both the estimates of the coupling and the estimates of the coupling variance are affected. While affected, one notes that when the Hilbert transform instantaneous phase estimator is consistent and when model mis-specification is sufficiently small, the Hilbert transform instantaneous phase estimator approaches the maximum likelihood instantaneous phase estimator for typically sized neuroscience data. In this situation, substitution of the Hilbert-type instantaneous phase estimator for the maximum likelihood instantaneous phase estimator in the score equations, Eq. (41), yields, upon extremization, the maximum likelihood estimate of $\boldsymbol{\beta}$ to acceptable accuracy.

Numerical computation of the maximum likelihood estimate of $\boldsymbol{\beta}$ is accomplished with a modified Newton–Raphson algorithm. Let the mth element of the error vector $\mathbf{e}^i$, on the ith iteration be,

$$(\mathbf{e}^{(i)}(\boldsymbol{\beta}^{(i)},\hat{\boldsymbol{\phi}}))_m = \frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\phi})}{\partial\beta_m}. \tag{48}$$

Here, $\hat{\boldsymbol{\phi}}$ is the Hilbert-type instantaneous phase estimator, Let the maximum likelihood estimate of $\boldsymbol{\beta}$ equal $\tilde{\boldsymbol{\beta}}$. Then, $\mathbf{e}^{(i)}(\tilde{\boldsymbol{\beta}},\hat{\boldsymbol{\phi}})$ approximates the zero vector, $\mathbf{0}$, for any iteration i. Note that the log-likelihood, Eq. (34), is undefined for λ_t equal to zero. Start the Newton–Raphson iterations with an initial guess of $\tilde{\boldsymbol{\beta}}$, set to

$$\tilde{\boldsymbol{\beta}}^{(0)} = [(|N(f_0)| + \epsilon)\ Re\{N(f_0)\}\ Im\{N(f_0)\}]^T, \tag{49}$$

where $N(f)$ is the discrete Fourier transform of dn_t evaluated at the frequency of interest, f_0. The operators, $Re\{\,\}$, $Im\{\,\}$ denote, respectively, taking the real and imaginary components. The constant, ϵ is small and positive. By initializing $\tilde{\boldsymbol{\beta}}$ according to Eq. (49), the log-likelihood and its derivatives are defined for all time-indices, t,

[10] $\mathbf{H}^T\mathbf{H}$ is diagonal when the instantaneous LFP phase is $\phi_t = 2\pi f_0 t\Delta$, where the frequency, f_0 is specified to be an integer multiple of the Rayleigh resolution; i.e. $f_0 = j/(N\Delta)$, for j integer.

on the first iteration. During the Newton–Raphson iterative procedure described below, λ_t may approach zero for some time-indices t. As previously mentioned, when this occurs the log-likelihood, Eq. (34), is undefined. The following strategy is employed. Only those time-indices where λ_t is greater than the small positive constant ϵ are retained in the Newton–Raphson procedure for any given iteration, i. Thus, with this restriction, the log-likelihood and its first two derivatives are well behaved. To obtain $\delta\boldsymbol{\beta}^{(i)}$, the change in $\boldsymbol{\beta}^{(i)}$, from one iteration to the next solve the following equation:

$$\mathbf{e}^{(i+1)} = \mathbf{e}^{(i)} + \mathbf{J}_{(i)}\delta\boldsymbol{\beta}^{(i)}, \tag{50}$$

$$\mathbf{e}^{(i+1)} = \mathbf{0}. \tag{51}$$

Here $\mathbf{J}_{(i)}$ is the Hessian matrix, at iteration i, specified by Eq. (45), and evaluated at $\lambda(\boldsymbol{\beta}^{(i)}, \hat{\boldsymbol{\phi}})$. Thus,

$$\delta\boldsymbol{\beta}^{(i)} = -\mathbf{J}_{(i)}^{-1}\mathbf{e}^{(i)}, \tag{52}$$

and

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} + \delta\boldsymbol{\beta}^{(i)}. \tag{53}$$

Iterations are continued until the elements of $\mathbf{e}^{(i)}$ are sufficiently close to zero. One notes that the effect of restricting this algorithm to work only with those time indices where λ_t is greater than ϵ is equivalent to discarding data. Thus, there is a potential for information loss, which is accompanied by the possibility of identifying an improved estimator. In practice, the above algorithm works well.

Appendix B. Asymptotic convergence

As mentioned in Section 2.1, standard theorems regarding the asymptotic convergence of the distribution of parameter estimators to the Gaussian distribution require the link, at the very least, to be everywhere differentiable (McCullagh and Nelder, 1999; Hogg et al., 2005; Pawitan, 2001). While the piece-wise linear link function used to specify Eq. (8) is not differentiable at the origin, due to the Weierstrass approximation theorem (Rudin, 1976), it can be uniformly approximated, on an interval $[a, b]$ containing rates of practical interest, as closely as desired by a polynomial function. Since polynomial functions are differentiable to arbitrarily high orders the regularity conditions required to ensure asymptotic normality of parameter maximum likelihood estimators holds for functions that are arbitrarily close to the piece-wise linear link function employed in this work.

A simulation is performed to study the convergence of the distributions of the maximum likelihood estimators to the normal distribution. A set of synthetic experiments is created; each consisting of a single trial. The results are presented in Fig. 15. For each trial, one second of measurements is synthesized. For both simulations, the coupling constant, β, linking the LFP to the rate is 100 Hz, and the LFP is drawn from the same autoregressive (AR) process used to specify the LFP in Section 4. The background rate of the neuron is increased from 0 Hz, to 300 Hz from the first row to the fourth row in both figures. For each trial and background rate configuration, 150 realizations of the experimental data are computed and for each realization maximum likelihood estimates of the parameters, β_0, β_c and β_s. These 150 estimates are then used to compute estimates of the probability density functions for the associated estimators, $\hat{\beta}_0$, $\hat{\beta}_c$ and $\hat{\beta}_s$. The red curves are Gaussian probability density functions specified to have a mean equal to the sample mean of the fifty estimates, and a variance specified by the observed Fisher information computed using Eq. (46). Fig. 15 indicates that the Gaussian approximation is good for all rates and is excellent for rates greater than or equal to 50 Hz. Since the non-linear link function is in a linear regime when the background rate exceeds 100 Hz, one expects in this regime estimator variance to increase with background rate; consistent with the Poisson distribution. This can be seen by inspecting the width of the probability density functions in the last two rows of Fig. 15. When the background rate is 0 Hz, the rate is a rectified cosine, and though bias is evident; the estimates are sufficiently accurate to usefully perform inference.

Appendix C. Effect of unmodelled sinusoids

In general, one expects neural activity to be influenced simultaneously by multiple sinusoids of varying frequencies. In this situation the model specified by Eqs. (8) and (9) is incorrect. This mis-specification can manifest in biased parameter estimates as well as biased estimates of parameter variance and covariance. The following two lemmas address this concern.

Lemma 1. *Unbiased estimator for oscillatory model with strictly positive rate.*

Let $\mathbf{c}(f)$ *and* $\mathbf{s}(f)$ *be two column vectors with their jth elements equal to* $\cos(2\pi f\Delta j)$ *and* $\sin(2\pi f\Delta j)$*, respectively. Here* Δ *is the bin size employed in Section 2. Let the jth Fourier frequency,* f_j*, equal* j/T *where T is the observation duration. Define the complete model matrix,* $\mathbf{H}_f$*, as,*

$$\mathbf{H}_f = \frac{1}{\sqrt{N}}\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{1} & \mathbf{c}(f_1) & \mathbf{s}(f_1) & \dots & \mathbf{c}(f_{N-1}) & \mathbf{s}(f_{N-1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Given the element-by-element restriction,

$$P(\mathbf{H}_f\boldsymbol{\beta}_f < \mathbf{0}) = 0, \tag{54}$$

there exists a spectral representation Priestly (1981) of a finite, discrete-time random process such that the stochastic intensity for the full model incorporating all sinusoids can be represented as the piece-wise linear function of a linear combination of random parameters, $\boldsymbol{\beta}_f$*:*

$$\lambda_t = \max(0, (\mathbf{H}_f\boldsymbol{\beta}_f)_t). \tag{55}$$

Eq. (55), when considered along with the restriction, Eq. (54), describes neural activity influenced by up to the maximal number of sinusoids permitted by Fourier theory when the background rate is sufficiently large. Then the expected intensity for all considered time-indices, $E\{\boldsymbol{\lambda}\}$*, is,*

$$\begin{aligned} E\{\boldsymbol{\lambda}\} &= E\{\boldsymbol{\lambda} \mid \mathbf{H}_f\boldsymbol{\beta}_f \geq \mathbf{0}\}P(\mathbf{H}_f\boldsymbol{\beta}_f \geq \mathbf{0}) + E\{\boldsymbol{\lambda} \mid \mathbf{H}_f\boldsymbol{\beta}_f < \mathbf{0}\}P(\mathbf{H}_f\boldsymbol{\beta}_f < \mathbf{0}), \\ &= E\{\boldsymbol{\lambda} \mid \mathbf{H}_f\boldsymbol{\beta}_f \geq \mathbf{0}\}, \\ &= \Delta\mathbf{H}_f E\{\boldsymbol{\beta}_f\}. \end{aligned} \tag{56}$$

Let the three-column restricted model matrix, $\mathbf{H}_r$ *equal to* $\mathbf{H}_r = (1/\sqrt{N})[\mathbf{1}\ \mathbf{c}(f_i)\ \mathbf{s}(f_i)]$*, where* f_i *is the Fourier frequency of interest. Let the triplet,* $\boldsymbol{\beta}_a$*, of elements of* $\boldsymbol{\beta}_f$*, consist of the pair of random parameters associated with the frequency of interest,* f_i*, and the random parameter associated with the first column of the model matrix, representing the constant, or zero frequency component. Further restrict every element of the product,* $\mathbf{H}_r\boldsymbol{\beta}_r$*, to be positive:*

$$P(\mathbf{H}_r\boldsymbol{\beta}_r < \mathbf{0}) = 0. \tag{57}$$

Then, the maximum-likelihood estimator, $\hat{\boldsymbol{\beta}}_r$*, computed with the restricted model matrix,* $\mathbf{H}_r$*, is unbiased.*

Proof. The maximum-likelihood estimator, $\hat{\boldsymbol{\beta}}_r$, zeros the score equation:

$$\mathbf{H}_r^T(\mathbf{dn} - \Delta\max(0,\ \mathbf{H}_r\hat{\boldsymbol{\beta}}_r)) = \mathbf{0}. \tag{58}$$

Taking the expectation of both sides of Eq. (58) yields,

$$E\{\mathbf{H}_r^T(\mathbf{dn} - \Delta\ \max(0,\ \mathbf{H}_r\hat{\boldsymbol{\beta}}_r))\} = \mathbf{H}_r^T[\mathbf{H}_f E\{\boldsymbol{\beta}_f\} - \mathbf{H}_r E\{\hat{\boldsymbol{\beta}}_r\}] = \mathbf{0}. \quad (59)$$

Due to orthogonality, Eq. (59) can be written:

$$\mathbf{H}_r^T\mathbf{H}_r(E\{\boldsymbol{\beta}_a\} - E\{\hat{\boldsymbol{\beta}}_r\}) = 0. \quad (60)$$

Because the null space of $\mathbf{H}_r$ contains the zero element only, Eq. (60) implies $\boldsymbol{\beta}_a$ equals $\hat{\boldsymbol{\beta}}_r$ in expectation. □

Lemma 1 states that the parameter estimators associated with the full and reduced models are identical when the model matrix is oscillatory and the rate is sufficiently high. When the rate is sufficiently high nonlinear action of the piece-wise linear link function is not in effect. One notes that the oscillatory model matrix is approximately attained when the instantaneous phase of the LFP is low dimensional. That is, when the product of the observation duration of the LFP multiplied by the width of the frequency interval of the band-pass filter used in the Hilbert transform based LFP phase estimator is small Slepian (1976). This result is intuitive, as in this situation the maximum likelihood estimate in the proposed model is directly related to the discrete Fourier transform of the counts **dn**, evaluated at the frequency of interest. For the testing methodology proposed in this work, the Fisher information associated with the full model restricted to the parameters of interest and the Fisher information associated with the reduced model should be equal to avoid biased scaling effects in the proposed hypothesis test.

Lemma 2. *Invariance of the Fisher Information to non-interesting sinusoids.*

With the definitions and restrictions specified in Lemma 1, the Fisher information, $\mathbf{I}_{\beta_a}$ associated with the fully specified model and the Fisher information for the reduced model, $\mathbf{I}_{\beta_r}$, are equal.

Proof. The Fisher information associated with the reduced model estimator, $\hat{\boldsymbol{\beta}}_r$, $\mathbf{I}_{\beta_r}$ is, by definition Scharf (1991) and Kay (1993),

$$\mathbf{I}_{\beta_r} = E\{\mathbf{s}(\boldsymbol{\beta}_r, \mathbf{dn})\,\mathbf{s}(\boldsymbol{\beta}_r, \mathbf{dn})^T\}, \quad (61)$$

where the score equation, $\mathbf{s}(\boldsymbol{\beta}_r, \mathbf{dn})$, is

$$\mathbf{s}(\boldsymbol{\beta}_r, \mathbf{dn}) = \mathbf{H}_r^T(\mathbf{dn} - \Delta \max(0, \mathbf{H}_r\boldsymbol{\beta}_r)). \quad (62)$$

Then,

$$\mathbf{I}_{\beta_r} = E\{\mathbf{T}_1^{(r)} + \mathbf{T}_2^{(r)} + \mathbf{T}_3^{(}r) + \mathbf{T}_4^{(r)}\}. \quad (63)$$

Equality between $\mathbf{I}_{\beta_r}$ and $\mathbf{I}_{\beta_a}$ is established when the matrices $E\{\mathbf{T}_j^{(r)}\}, j = 1, 2, 3, 4$ associated with the reduced model are shown to be equivalent to those matrices, $E\{\mathbf{T}_j^{(a)}\}$ associated with the Fisher information of the full model restricted to the parameters of interest. Consider,

$$\begin{aligned}
E\{\mathbf{T}_1^{(r)}\} &= \mathbf{H}_r^T E\{\mathbf{dn}\,\mathbf{dn}^T\}\mathbf{H}_r,\\
E\{\mathbf{T}_1^{(r)}\} &= \mathbf{H}_r^T\mathbf{H}_f E\{\boldsymbol{\beta}_f\boldsymbol{\beta}_f^T\}\mathbf{H}_f^T\mathbf{H}_r,\\
E\{\mathbf{T}_1^{(r)}\} &= \mathbf{H}_r^T[\mathbf{H}_r\,\mathbf{0}]E\{\boldsymbol{\beta}_f\boldsymbol{\beta}_f^T\}[\mathbf{H}_r\,\mathbf{0}]^T\mathbf{H}_r,\\
E\{\mathbf{T}_1^{(r)}\} &= \mathbf{H}_r^T\mathbf{H}_r E\{\boldsymbol{\beta}_a\boldsymbol{\beta}_a^T\}\mathbf{H}_r^T\mathbf{H}_r,\\
E\{\mathbf{T}_1^{(r)}\} &= E\{\mathbf{T}_1^{(a)}\}.
\end{aligned} \quad (64)$$

Where orthonormality is used to proceed from line 4 to line 5. Similar arguments hold for $j = 2, 3, 4$ and the equality of the Fisher information matrices is established. □

Hence, there is no model mis-specification due to unmodelled sinusoids when the rate is sufficiently high such that the nonlinear function relating the linear combination of covariates to the conditional intensity is always in the linear regime. At lower rates one expects an effect due to unmodelled sinusoids. The form of this effect is best explored through standard, statistical model selection procedures.

Appendix D. Approximate modulation confidence interval

By Taylor expansion (Casella and Berger, 2001, p. 242), the variance, $var\{\hat{\rho}\}$ of the modulation estimator, $\hat{\rho}$ can be approximated as,

$$var\{\hat{\rho}\} = \frac{\hat{\beta}_c^2\ var\{\hat{\beta}_c\} + \hat{\beta}_s^2\ var\{\hat{\beta}_s\}}{\hat{\rho}^2}. \quad (65)$$

Here, *var* denotes variance, and ˆ over a symbol denotes an estimator of the theoretical quantity. In this work, confidence intervals are constructed from Eq. (65) assuming normality. Relevant estimates replace the estimators in Eq. (65) to obtain an estimate of the variance of $\hat{\rho}$. This estimate is used to compute confidence intervals. One notes that if one does not mind being conservative, greater accuracy can be attained by employing the Chebyshev inequality to obtain the confidence interval. In this case, an assumption of normality is not required.

References

Akaike H. A new look at the statistical model identification. IEEE Trans Automat Control 1974;19(6):716–23.

Amjad AM, Halliday DM, Rosenberg JR, Conway BA. An extended difference of coherence test for comparing and combining several independent coherence estimates: theory and application to the study of motor units and physiological tremor. J Neurosci Methods 1997;73(2/3):69–79.

Bollimunta A, Chen Y, Schroeder CE, Ding M. Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques. J Neurosci 2008;28(40):9976–88.

Box GEP, Draper NR. Empirical model-building and response surfaces. US, Canada: Wiley; 1987.

Brillinger DR. The identification of point process systems. Ann Probabil 1975;3(6):909–24.

Brillinger DR. Time series: data analysis and theory. San Francisco: Society for Industrial and Applied Mathematics; 2001.

Bruns A. Fourier-, hilbert- and wavelet-based signal analysis: are they really different approaches? J Neurosci Methods 2004;137(2):321–32.

Bruns A, Eckhorn R. Task-related coupling from high- to low-frequency signals among visual cortical areas in human subdural recordings. Int J Psychophysiol 2004;51(2):97–116.

Bullock TH, Mcclune MC, Achimowicz JZ, Iragui-Madoz VJ, Duckrow RB, Spencer SS. Temporal fluctuations in coherence of brain waves. Proc Natl Acad Sci U S A 1995;92(25):11568–72.

Casella G, Berger RL. Statistical inference. 2nd edition Duxbery Press; 2001.

Chalk M, Herrero JL, Gieselmann MA, Delicato LS, Gotthardt S, Thiele A. Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in v1. Neuron 2010;66(1):114–25.

Daley DJ, Vere-Jones D. An introduction to the theory of point processes. Springer Series in Statistics; 2003.

DeCoteau WE, Thorn C, Gibson DJ, Courtemanche R, Mitra P, Kubota Y, et al. Learning-related coordination of striatal and hippocampal theta rhythms during acquisition of a procedural maze task. Proc Natl Acad Sci U S A 2007a;104(13):5644–9.

DeCoteau WE, Thorn C, Gibson DJ, Courtemanche R, Mitra P, Kubota Y, et al. Oscillations of local field potentials in the rat dorsal striatum during spontaneous and instructed behaviors. J Neurophysiol 2007b;97(5):3800–5.

Feller W. An introduction to probability theory, vol. II. New York: Wiley and Sons; 1966.

Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. Science 2001;291(5508):1560–3.

Fries P, Womelsdorf T, Oostenveld R, Desimone R. The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area v4. J Neurosci 2008;28(18):4823–35.

Gregoriou G, Gotts SJ, Zhou H, Desimone R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science 2009;324(5931):1207–10.

Halliday DM, Rosenberg JR, Amjad AM, Breeze P, Conway BA, Farmer SF. A framework for the analysis of mixed time series/point process data – theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. Prog Biophys Mol Biol 1995;64(2/3):237–78.

Hogg RV, McKean JW, Craig AT. Introduction to mathematical statistics. 6th ed. Pearson Prentice Hall; 2005.

Jarvis M, Mitra P. Sampling properties of the spectrum and coherency of sequences of action potentials. Neural Comput 2001;13(4):717.

Jutras MJ, Fries P, Buffalo EA. Gamma-band synchronization in the macaque hippocampus and memory formation. J Neurosci 2009;29(40):12521–31.

Kay SM. Fundamentals of statistical signal processing: estimation theory. Prentice Hall; 1993.
Koch C. Biophysics of computation. Oxford University Press; 1999.
Kristeva R, Patino L, Omlor W. Beta-range cortical motor spectral power and corticomuscular coherence as a mechanism for effective corticospinal interaction during steady-state motor output. Neuroimage 2007;36(3): 785–92.
Lepage KQ, Kramer MA, Eden UT. The dependence of spike field coherence on expected intensity. Neural Comput 2011;23(9):2209–41.
MacDonald CJ, Lepage KQ, Eden UT, Eichenbaum H. Hippocampal time cells bridge the gap in memory for discontiguous events. Neuron 2011;71(4):737–49.
Mardia K, Jupp P. Directional statistics. Wiley; 2000.
McCullagh P, Nelder JA. Generalized linear models. 2nd ed. Chapman & Hall/CRC; 1999.
Mitchell JF, Sundberg KA, Reynolds JH. Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. Neuron 2009;63(6):879–88.
Mitra P, Bokil H. Observed brain dynamics. Oxford University Press; 2008.
Montgomery SM, Buzsáki G. Gamma oscillations dynamically couple hippocampal ca3 and ca1 regions during memory task performance. Proc Natl Acad Sci U S A 2007;104(36):14495–500.
Pawitan Y. In all likelihood: statistical modelling and inference using likelihood. 1st ed. Oxford: Clarendon Press; 2001.
Pesaran B, Nelson MJ, Andersen RA. Free choice activates a decision circuit between frontal and parietal cortex. Nature 2008;453(7193):406–9.
Priestly MB. Spectral analysis and time series. Elsevier Academic Press; 1981.
Rice S. Mathematical analysis of random noise-conclusion. Bell Systems Tech J 1945;24:46–156.
Rosenberg JR, Halliday DM, Breeze P, Conway BA. Identification of patterns of neuronal connectivity – partial spectra, partial coherence, and neuronal interactions. J Neurosci Methods 1998;83(1):57–72.
Rudin W. Principles of mathematical analysis. 3rd ed. McGraw-Hill; 1976.
Sanger TD. Theoretical considerations for the analysis of population coding in motor cortex. Neural Comput 1994;6(1):29–37.
Scharf LL. Statistical signal processing: detection, estimation, and time series analysis. Addison-Wesley Publishing Company, Inc; 1991.
Sirota A, Montgomery S, Fujisawa S, Isomura Y, Zugaro M, Buzsáki G. Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. Neuron 2008;60(4):683–97.
Slepian D. On bandwidth. Proc IEEE 1976;64(3):292–300.
Towle VL, Carder RK, Khorasani L, Lindberg D. Electrocorticographic coherence patterns. J Clin Neurophysiol 1999;16(6):528–47.
Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. J Neurophysiol 2005;93(2):1074–89.
Vinck M, Battaglia F, Womelsdorf T, Pennartz C. Improved measures of phase-coupling between spikes and the local field potential. J Comput Neurosci 2011;55(4):1–23.
Vinck M, Van Wingerden M, Womelsdorf T, Fries P, Pennartz C. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage 2010;51(1):112–22.
Witham CL, Wang M, Baker SN. Cells in somatosensory areas show synchrony with beta oscillations in monkey motor cortex. Eur J Neurosci 2007;26(9): 2677–86.
Womelsdorf T, Fries P, Mitra P, Desimone R. Gamma-band synchronization in the visual cortex predicts speed of change detection. Nature 2006;439(7077): 733–6.
Zaveri HP, Williams WJ, Sackellares JC, Beydoun A, Duckrow RB, Spencer SS. Measuring the coherence of intracranial electroencephalograms. Clin Neurophysiol 1999;110(10):1717–25.